Gaussian-Basis Monte Carlo Method for Numerical Study on Ground States of Itinerant and Strongly Correlated Electron Systems

Takeshi Aimi[1] and Masatoshi Imada[2]

Department of Physics[1] *and Department of Applied Physics*[2]*, University of Tokyo*
7-3-1 Hongo, Bunkyo-ku, Tokyo, Japan

We examine Gaussian-basis Monte Carlo method (GBMC) introduced by Corney and Drummond. This method is based on an expansion of the density-matrix operator $\hat{\rho}$ by means of the coherent Gaussian-type operator basis $\hat{\Lambda}$ and does not suffer from the minus sign problem. The original method, however, often fails in reproducing the true ground state and causes systematic errors of calculated physical quantities because the samples are often trapped in some metastable or symmetry broken states. To overcome this difficulty, we combine the quantum-number projection scheme proposed by Assaad, Werner, Corboz, Gull and Troyer in conjunction with the importance sampling of the original GBMC method. This improvement allows us to carry out the importance sampling in the quantum-number-projected phase-space. Some comparisons with the previous quantum-number projection scheme indicate that, in our method, the convergence with the ground state is accelerated, which makes it possible to extend the applicability and widen the range of tractable parameters in the GBMC method. The present scheme offers an efficient practical way of computation for strongly correlated electron systems beyond the range of system sizes, interaction strengths and lattice structures tractable by other computational methods such as the quantum Monte Carlo method.

1. Introduction

Ground state properties of strongly correlated electron systems are challenging subjects in condensed matter physics. From the numerical point of view, there exist many numerical algorithms, such as the exact diagonalization method, the auxiliary-field quantum Monte Carlo (AFQMC) method,[1–4] the density matrix renormalization group (DMRG) method[5] and the path-integral renormalization group (PIRG) method.[6–11] Although the exact diagonalization of the Hamiltonian matrix gives accurate results, the tractable system size is severely limited. The AFQMC method can treat larger systems and has been applied to various correlated systems. In some systems such as doped Mott insulators and the Mott insulators with geometrical frustration effects, however, the AFQMC method often suffers from the negative sign problem which causes large statistical errors by the cancellation of positive and negative Monte Carlo samples. The DMRG method offers practically exact results without suffering from the negative sign problem. However, tractable lattice systems are restricted to one-dimensional configurations owing to the spatial renormalization process. The PIRG method is a powerful sign-free numerical technique for correlated electron systems and has been applied to various systems beyond the tractable range of the above numerical methods. A main practical limitation in the PIRG method comes from the extrapolation procedure to reach the results for the full Hilbert space. The truncation error depends on the system size as well as on the interaction strength.

Gaussian-basis Monte Carlo (GBMC) method has

been proposed as an alternative quantum Monte Carlo method which does not involve any sign problem.[12, 13] This method is based on a representation of the density-matrix operator $\hat{\rho}$ by making use of the non-Hermitian Gaussian-type operator basis $\hat{\Lambda}$. The Gaussian representation is a natural generalization of the positive-P phase-space method which is often used in the area of quantum optics.[14, 15] As well as classical phase-space variables like $(\boldsymbol{x}, \boldsymbol{p})$, the Gaussian-basis representation utilizes the one-particle "Green's function" $n_{ij} = \mathrm{Tr}\left[\hat{c}_i^\dagger \hat{c}_j \hat{\Lambda}\right]$, $m_{ij} = \mathrm{Tr}\left[\hat{c}_i \hat{c}_j \hat{\Lambda}\right]$, $m_{ij}^+ = \mathrm{Tr}\left[\hat{c}_i^\dagger \hat{c}_j^\dagger \hat{\Lambda}\right]$ and the stochastic weight Ω as the phase-space variables, where $\hat{c}_i^\dagger$ ($\hat{c}_i$) is a Fermion creation (annihilation) operator of the i-th mode. In this method, we solve a Fokker-Planck equation with respect to the phase-space variables $\underline{\lambda} = (\Omega, \boldsymbol{n}, \boldsymbol{m}, \boldsymbol{m}^+)$, which is constructed by a mapping from an operator Liouville equation of $\hat{\rho}$. One of the phase-space variables Ω works as a weight of the importance sampling and for any two-body Hamiltonian, Ω remains positive definite. Thus there exists no explicit manifestation of the negative sign problem. However, in many parameter regions, especially in the low-temperature region, the numerical results obtained by the GBMC method often show systematic errors.[16, 17] It will turn out below that this deviation is concerned with the "spontaneous symmetry breaking".

Assaad *et al.* have used a quantum-number projection scheme to overcome the deviations in the low-temperature region.[16, 17] They have proposed to project the density matrix onto given quantum numbers of the ground state after the sampling is completed and have reproduced accurate ground states in some parameter

regions. In their method, however, the convergence with the ground state becomes slower with the increase of the interaction strength, which determines the practical limitations.

In this study, we combine the quantum-number projection scheme concurrently with the importance sampling of the original GBMC method to reflect the amount of the overlap with the projected sector in the sampling weight. This allows us to perform the importance sampling with respect to the projected distribution. The crucial point is that the efficiency of the importance sampling is improved because the sampling weight reflects not only the energy but also the overlap with the projected state, *i.e.*, the overlap with the state which retains the same quantum numbers with the ground state. Thanks to the efficient sampling, the tractable parameter region becomes wider than the previous method reported by Assaad *et al.*[16, 17] Moreover, our method allows us to analyze the projected distribution directly. By using this advantage, the relation between the numerical convergence and the behavior of the projected distribution is also reported. We show benchmark analysis up to 10×10 lattices on the square lattice as well as up to the relative interaction strength $U/t = 15$ for the Hubbard model, which indicate the applicability and efficiency of the present method.

The organization of this paper is as follows. Section 2 gives an introduction of the GBMC method for the general Fermion systems by following the formulation in Refs. 12, 13, which we supply for the self-contained description. In §3, we explain implementations of the Monte Carlo procedure for the Hubbard model. The improvements of the GBMC method by quantum-number projections are shown in §4. In the last part of §4, we discuss the practical limitation in the applicability of the GBMC method. Section 5 is devoted to summary and discussions.

2. Gaussian-Basis Monte Carlo Method

Gaussian-basis Monte Carlo (GBMC) method is a numerical method which makes use of the mapping between operator equations of motion and stochastic evolution equations of generalized phase-space.[12, 13] In order to make an exact mapping, we introduce a complete set of Gaussian-type operators $\hat{\Lambda}$, which is typically non-Hermitian. This basis set allows us to expand any physical density-matrix operator $\hat{\rho}(\tau)$ in terms of the phase-space variables $\underline{\lambda}$ as

$$\hat{\rho}(\tau) = \int d\underline{\lambda} P(\underline{\lambda}, \tau)\hat{\Lambda}(\underline{\lambda}), \qquad (1)$$

where τ is real or imaginary time, $P(\underline{\lambda}, \tau)$ is the expansion coefficient, and $d\underline{\lambda}$ is the integration measure. Here, $P(\underline{\lambda}, \tau)$ can always be chosen positive as in Appendix B and hence can be regarded as a probability distribution. In the GBMC method, it is the distribution $P(\underline{\lambda}, \tau)$ that is sampled stochastically.

2.1 General Gaussian basis
2.1.1 Notation

Before defining the Gaussian basis, we summarize the notation which will be used. Consider a M-mode Fermionic system characterized by the creation and annihilation operators $\hat{c}_j^\dagger$ and $\hat{c}_j$, with anticommutation relations

$$[\hat{c}_i, \hat{c}_j^\dagger]_+ = \delta_{i,j}, \quad [\hat{c}_i, \hat{c}_j]_+ = [\hat{c}_i^\dagger, \hat{c}_j^\dagger]_+ = 0, \qquad (2)$$

where $i, j = 1, \ldots, M$. We define a M-mode column vector of the annihilation operators and its Hermitian conjugate row vector as $\hat{\boldsymbol{c}}$ and $\hat{\boldsymbol{c}}^\dagger$, respectively. In order to define the general Gaussian basis in a compact form, we introduce an extended-vector notation

$$\underline{\hat{c}} = \left(\begin{array}{c} \hat{\boldsymbol{c}} \\ \hat{\boldsymbol{c}}^{\dagger T} \end{array} \right), \quad \underline{\hat{c}}^\dagger = \left(\hat{\boldsymbol{c}}^\dagger, \hat{\boldsymbol{c}}^T \right). \qquad (3)$$

Throughout the paper, we use the bold-type notation for M-mode vectors or matrices and the underline notation for $2M$-mode extended vectors or matrices.

For products of operators, we define a normal and an antinormal ordering operators denoted by : : and { }, respectively. The normal ordering operator : : reorders so that all the creation operators are put to the left of the annihilation operators, *e.g.*, $: \hat{c}_i \hat{c}_j^\dagger := -\hat{c}_j^\dagger \hat{c}_i$. Similarly, the antinormal ordering operator { } reorders so that all the annihilation operators are put to the left of the creation operators, *e.g.*, $\{\hat{c}_j^\dagger \hat{c}_i\} = -\hat{c}_i \hat{c}_j^\dagger$. More generally, in the case of a nested product, the outer ordering operator does not reorder the inner one, *e.g.*, $\{: \hat{c}_k \hat{c}_j^\dagger : \hat{c}_i\} = \hat{c}_i : \hat{c}_k \hat{c}_j^\dagger := -\hat{c}_i \hat{c}_j^\dagger \hat{c}_k$. The sign changes are necessary because of the anticommuting nature of the Fermion operators.

2.1.2 Definition of the Gaussian basis

Using the above notation, a general Gaussian operator is defined as :

$$\hat{\Lambda}(\underline{\lambda}) = \Omega \text{Pf}[\underline{\sigma}_A] : \exp\left[\underline{\hat{c}}^\dagger (\underline{I} - \underline{\sigma}^{-1}/2)\underline{\hat{c}} \right] :, \qquad (4)$$

where $\underline{I}$ is an extended unit matrix:

$$\underline{I} = \left[\begin{array}{cc} -\boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I} \end{array} \right], \qquad (5)$$

$\underline{\sigma}$ is an extended covariance:

$$\underline{\sigma} = \left[\begin{array}{cc} \boldsymbol{n}^{\mathrm{T}} - \boldsymbol{I} & \boldsymbol{m} \\ \boldsymbol{m}^+ & \boldsymbol{I} - \boldsymbol{n} \end{array} \right], \qquad (6)$$

and the vector parameter $\underline{\lambda}$ is defined as:

$$\underline{\lambda} = (\Omega, \boldsymbol{n}, \boldsymbol{m}, \boldsymbol{m}^+). \qquad (7)$$

Here, $\boldsymbol{n}$ is a $M \times M$ matrix which corresponds to normal Green's function, while $\boldsymbol{m}$ and $\boldsymbol{m}^+$ are two independent $M \times M$ antisymmetric matrices which correspond to anomalous Green's functions. Pfaffian of the antisymmetrized covariance $\underline{\sigma}_A$ appears so as to satisfy the normalization condition

$$\text{Tr}[\hat{\Lambda}(\underline{\lambda})] = \Omega. \qquad (8)$$

Here, $\underline{\sigma}_A$ is constructed by moving each row in the lower half rows every after the row with the same indices in the upper half, and moving each column in the right half

columns every before the column with the same indices in the left half, *i.e.*,

$$
\begin{bmatrix} \mathbf{a} & \mathbf{b} \\ \mathbf{c} & \mathbf{d} \end{bmatrix}_A = \begin{bmatrix} b_{11} & a_{11} & \cdots & b_{1M} & a_{1M} \\ d_{11} & c_{11} & \cdots & d_{1M} & c_{1M} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ b_{M1} & a_{M1} & \cdots & b_{MM} & a_{MM} \\ d_{M1} & c_{M1} & \cdots & d_{MM} & c_{MM} \end{bmatrix}. \quad (9)
$$

The general Gaussian operator $\hat{\Lambda}(\underline{\lambda})$ itself may correspond to a density-matrix operator under the conditions that $\boldsymbol{n}^\dagger = \boldsymbol{n}$, $\boldsymbol{m}^+ = \boldsymbol{m}^\dagger$ and the eigenvalues of the matrix $\boldsymbol{n}$ lie in the interval $[0, 1]$. However, we do not restrict the Gaussian operator to such conditions and it allows the Gaussian basis to be an overcomplete set which can expand any physical density-matrix operator.

2.1.3 Properties of Gaussian basis

From the above definition, the general Gaussian operators satisfy some important identities.[12, 13] First, the Gaussian operator operated by the Fermion operator $\hat{c}$ and $\hat{c}^\dagger$ can be associated with differentiations of the Gaussian operators with respect to their parameters:

$$
\hat{\Lambda} = \Omega \frac{\partial}{\partial \Omega} \hat{\Lambda}, \quad (10)
$$

$$
: \underline{\hat{c}}\,\underline{\hat{c}}^\dagger \hat{\Lambda} : = \underline{\underline{\sigma}}\hat{\Lambda} - \underline{\underline{\sigma}}\frac{\partial \hat{\Lambda}}{\partial \underline{\underline{\sigma}}}\underline{\underline{\sigma}}, \quad (11)
$$

$$
\left\{ \underline{\hat{c}} : \underline{\hat{c}}^\dagger \hat{\Lambda} : \right\} = -\underline{\underline{\sigma}}\hat{\Lambda} + (\underline{\underline{\sigma}} - \underline{\underline{I}})\frac{\partial \hat{\Lambda}}{\partial \underline{\underline{\sigma}}}\underline{\underline{\sigma}}. \quad (12)
$$

Second, the traces of ladder operators with respect to the Gaussian operator can be analytically taken by using Grassmann coherent states:

$$
\mathrm{Tr}\left[\hat{c}_i^\dagger \hat{c}_j \hat{\Lambda} \right] = \Omega n_{ij}, \quad (13)
$$

$$
\mathrm{Tr}\left[\hat{c}_i \hat{c}_j \hat{\Lambda} \right] = \Omega m_{ij}, \quad \mathrm{Tr}\left[\hat{c}_i^\dagger \hat{c}_j^\dagger \hat{\Lambda} \right] = \Omega m_{ij}^+. \quad (14)
$$

2.2 Time evolution

The time evolution of a density-matrix operator is determined by the Liouville equation

$$
\frac{d\hat{\rho}(t)}{dt} = \hat{L}\hat{\rho}(t), \quad (15)
$$

where $\hat{L}$ is a Liouville superoperator. For a real time evolution, the superoperator is given by the commutator with the Hamiltonian:

$$
\hat{L}\hat{\rho}(t) \equiv -\frac{i}{\hbar}\left[\hat{H}, \hat{\rho}(t) \right]. \quad (16)
$$

In the case of calculating the equilibrium state at $\tau = 1/k_B T$, the imaginary time evolution of a density-matrix operator is determined by the equation

$$
\frac{d\hat{\rho}(\tau)}{d\tau} = -\frac{1}{2}\left[\hat{H}, \hat{\rho}(\tau) \right]_+. \quad (17)
$$

Therefore the superoperator is given by the anticommutator with the Hamiltonian:

$$
\hat{L}\hat{\rho}(\tau) \equiv -\frac{1}{2}\left[\hat{H}, \hat{\rho}(\tau) \right]_+ \quad (18)
$$

To construct a mapping, we first substitute the expansion in Eq. (1) into the Liouville equation (15) to get

$$
\int \frac{dP(\underline{\lambda}, t)}{dt}\hat{\Lambda}(\underline{\lambda})d\underline{\lambda} = \int P(\underline{\lambda}, t)\hat{L}\hat{\Lambda}(\underline{\lambda})d\underline{\lambda}. \quad (19)
$$

Second, using the differential properties in Eqs.(10-12), one can transform the superoperator $\hat{L}\hat{\Lambda}(\underline{\lambda})$ into a differential operator $\mathcal{L}\hat{\Lambda}(\underline{\lambda})$. We next apply partial integration to get, provided that boundary terms vanish,

$$
\int \hat{\Lambda}(\underline{\lambda})\frac{dP(\underline{\lambda}, t)}{dt}d\underline{\lambda} = \int \hat{\Lambda}(\underline{\lambda})\mathcal{L}'P(\underline{\lambda}, t)d\underline{\lambda}, \quad (20)
$$

where $\mathcal{L}'$ is reordered form of $\mathcal{L}$. Note that from Eq. (10-11) $\mathcal{L}'$ contains derivatives only up to the second order for any two-body Hamiltonian. As a sufficient solution for Eq. (20), a Fokker-Planck equation of Ito type is obtained:

$$
\frac{d}{dt}P(\underline{\lambda}, t) = \mathcal{L}'P(\underline{\lambda}, t)
$$

$$
= \left[-\sum_i \frac{\partial}{\partial \lambda_i}A_i(\underline{\lambda}) \right.
$$

$$
\left. + \frac{1}{2}\sum_{i,j} \frac{\partial}{\partial \lambda_i}\frac{\partial}{\partial \lambda_j}D_{ij}(\underline{\lambda}) \right] P(\underline{\lambda}, t). \quad (21)
$$

The imaginary-time evolution equation of the density-matrix operator then boils down to the Fokker-Planck equation, which is in practice solved by integrating numerically the corresponding stochastic differential equations (SDE).

3. Gaussian Representation for Hubbard Model

3.1 Mapping

We consider the following Hubbard Hamiltonian.

$$
\hat{H} = \sum_{i,j,\sigma}^N t_{ij}\hat{c}_{i\sigma}^\dagger \hat{c}_{j\sigma} + U\sum_i^N \hat{c}_{i\uparrow}^\dagger \hat{c}_{i\uparrow}\hat{c}_{i\downarrow}^\dagger \hat{c}_{i\downarrow}, \quad (22)
$$

$$
t_{ij} = \begin{cases} -\mu & \text{for } i = j, \\ -t & \text{for } (i, j) \text{ being a nearest-neighbor pair,} \\ 0 & \text{otherwise,} \end{cases} \quad (23)
$$

where i and j represent the lattice points, $\hat{c}_{i\sigma}^\dagger (\hat{c}_{i\sigma})$ the creation (annihilation) operator of an electron with spin σ on the i-th site, t_{ij} the transfer integral between the i-th site and the j-th site, U the on-site Coulomb interaction, μ the chemical potential and N the number of the lattice sites. Although we treat only this simplest Hubbard model, the formation can easily be extended to a more general form including transfers for further-site pairs and/or intersite Coulomb interactions.

Although Eq. (22) is a standard representation of the Hubbard Hamiltonian, it is necessary that the sign of the interaction term is negative so that the diffusion matrix $D(\underline{\lambda})$ in Eq. (21) being positive definite.[12, 13] Thus, we

transform the Hamiltonian as follows:[16, 17)]

$$\hat{H} = \hat{\boldsymbol{c}}^\dagger \boldsymbol{T} \hat{\boldsymbol{c}} - \frac{U}{2} \sum_i : (\hat{\boldsymbol{c}}_i^\dagger \sigma^z \hat{\boldsymbol{c}}_i)^2 : \tag{24}$$

$$= \sum_{x,y=1}^{2N} T_{xy} \hat{n}_{xy} + \frac{U}{2} \sum_{i=1}^N \sum_{\substack{\eta,\eta' \\ \sigma,\sigma'}} \left[\delta_{\sigma\eta'} \hat{n}_{(i\eta),(i\sigma')} \right.$$

$$\left. - \hat{n}_{(i\eta),(i\eta')} \hat{n}_{(i\sigma),(i\sigma')} \right] \sigma^z_{\eta\eta'} \sigma^z_{\sigma\sigma'}, \tag{25}$$

where $\hat{n}_{xy} = \hat{c}_x^\dagger \hat{c}_y$ and suffices x and y denote both the coordinates of site and spin, i.e., $x = (i,\sigma)$. The vector operators $\hat{\boldsymbol{c}}^\dagger$ and $\hat{\boldsymbol{c}}_i^\dagger$ are defined as

$$\hat{\boldsymbol{c}}^\dagger = (\hat{c}_{1\uparrow}^\dagger, \hat{c}_{2\uparrow}^\dagger, \cdots, \hat{c}_{N\uparrow}^\dagger, \hat{c}_{1\downarrow}^\dagger, \cdots, \hat{c}_{N\downarrow}^\dagger), \tag{26}$$

$$\hat{\boldsymbol{c}}_i^\dagger = (\hat{c}_{i\uparrow}^\dagger, \hat{c}_{i\downarrow}^\dagger). \tag{27}$$

The $2N \times 2N$ extended hopping matrix $\boldsymbol{T}$ is defined as $T_{i,j} = T_{i+N,j+N} = t_{i,j}$ and $T_{i,j+N} = T_{i+N,j} = 0$, where $i, j = 1, \cdots, N$. The matrix σ^z denotes the z component of the Pauli matrix. Since the Hubbard model conserves the total particle number, we use the number-conserving subset of the general Gaussian operator to expand the density-matrix operator:

$$\hat{\Lambda}(\Omega, \boldsymbol{n}) = \Omega \det(\boldsymbol{I} - \boldsymbol{n}) : e^{-\hat{\boldsymbol{c}}^\dagger [2\boldsymbol{I} + (\boldsymbol{n}^T - \boldsymbol{I})^{-1}] \hat{\boldsymbol{c}}} :, \tag{28}$$

where $\boldsymbol{n}$ is a $2N \times 2N$ matrix.

The Gaussian operator consists of an overcomplete set and it can expand any physical density-matrix operator with positive coefficients. In the following sections, we express the parameters of the Gaussian operator as $\underline{\lambda} = (\Omega, \boldsymbol{n})$. Similarly to the case of the general Gaussian operators in Eqs.(10-12), the number-conserving Gaussian satisfies the differential identities:

$$\hat{\Lambda} = \Omega \frac{\partial}{\partial \Omega} \hat{\Lambda}, \tag{29}$$

$$\hat{n}_{xy} \hat{\Lambda} = n_{xy} \hat{\Lambda} + (\delta_{xw} - n_{xw}) n_{zy} \frac{\partial \hat{\Lambda}}{\partial n_{zw}}, \tag{30}$$

$$\hat{\Lambda} \hat{n}_{xy} = n_{xy} \hat{\Lambda} + n_{xw} (\delta_{zy} - n_{zy}) \frac{\partial \hat{\Lambda}}{\partial n_{zw}}. \tag{31}$$

The trace of the Gaussian operator itself is $\text{Tr}[\hat{\Lambda}] = \Omega$ and the trace of any ladder operators can be calculated by Wick's theorem. For instance, we obtain

$$\text{Tr}[\hat{\Lambda} \hat{c}_x^\dagger \hat{c}_y] = \Omega n_{xy}, \tag{32}$$

$$\text{Tr}[\hat{\Lambda} \hat{c}_x^\dagger \hat{c}_y \hat{c}_w^\dagger \hat{c}_z] = \Omega \left[n_{xy} n_{wz} + n_{xz} (\delta_{wy} - n_{wy}) \right]. \tag{33}$$

To obtain the ground state of the system, one may consider the imaginary-time evolution of the density-matrix operator

$$\frac{\partial \hat{\rho}}{\partial \tau} = -\frac{1}{2} \left[\hat{H}, \hat{\rho} \right]_+, \qquad \tau = 1/k_B T. \tag{34}$$

In the GBMC method, instead of solving the Liouville equation above, one solves generalized Langevin equations by making use of the mapping between the Liouville equation and the stochastic equations. To this

end, we expand the density-matrix operator as $\hat{\rho} = \int d\underline{\lambda} P(\underline{\lambda}, \tau) \hat{\Lambda}(\underline{\lambda})$. Then the Liouville equation becomes

$$\int d\underline{\lambda} \hat{\Lambda}(\underline{\lambda}) \frac{\partial P(\underline{\lambda}, \tau)}{\partial \tau} = \int d\underline{\lambda} P(\underline{\lambda}, \tau) \left\{ -\frac{1}{2} \left[\hat{H}, \hat{\Lambda}(\underline{\lambda}) \right]_+ \right\}. \tag{35}$$

The differential identities of the Gaussian operator enable us to transform $-\frac{1}{2} \left[\hat{H}, \hat{\Lambda}(\underline{\lambda}) \right]_+$ into a differential form:

$$-\frac{1}{2} \left[\hat{H}, \hat{\Lambda} \right]_+ = \left[-\Omega H(\boldsymbol{n}) \frac{\partial}{\partial \Omega} - \sum_{x,y} A_{xy} \frac{\partial}{\partial n_{xy}} \right.$$

$$+ \frac{1}{2} \sum_i \sum_{\substack{x,y \\ w,z}} \left(B_{xy}^{(i)} B_{wz}^{(i)} \frac{\partial^2}{\partial n_{xy} \partial n_{wz}} \right.$$

$$\left. \left. + C_{xy}^{(i)} C_{wz}^{(i)} \frac{\partial^2}{\partial n_{xy} \partial n_{wz}} \right) \right] \hat{\Lambda}, \tag{36}$$

where

$$H(\boldsymbol{n}) = \text{Tr}[\hat{\Lambda}(\boldsymbol{n}) \hat{H}] / \text{Tr}[\hat{\Lambda}(\boldsymbol{n})], \tag{37}$$

$$\boldsymbol{A} = \frac{1}{2} \boldsymbol{n}(\boldsymbol{T} - U\boldsymbol{M})(\boldsymbol{I} - \boldsymbol{n})$$

$$+ \frac{1}{2}(\boldsymbol{I} - \boldsymbol{n})(\boldsymbol{T} - U\boldsymbol{M})\boldsymbol{n}, \tag{38}$$

$$B_{xy}^{(i)} = \sqrt{\frac{U}{2}} \sum_{\sigma,\sigma'} \sigma^z_{\sigma\sigma'} n_{x,(i\sigma')} (\delta_{(i\sigma),y} - n_{(i\sigma),y}), \tag{39}$$

$$C_{xy}^{(i)} = \sqrt{\frac{U}{2}} \sum_{\sigma,\sigma'} \sigma^z_{\sigma\sigma'} (\delta_{x,(i\sigma')} - n_{x,(i\sigma')}) n_{(i\sigma),y}, \tag{40}$$

$$M_{(i\sigma),(j\sigma')} = \delta_{ij} \sum_{\eta,\eta'} n_{(i\eta),(i\eta')} (\sigma^z_{\sigma\sigma'} \sigma^z_{\eta\eta'} - \sigma^z_{\sigma\eta'} \sigma^z_{\eta\sigma'}). \tag{41}$$

Partial integration, under the assumption that boundary terms vanish, yields the Fokker-Planck equation for the probability distribution $P(\underline{\lambda}, \tau)$:

$$\frac{\partial P(\underline{\lambda}, \tau)}{\partial \tau} = \left[\frac{\partial}{\partial \Omega} \Omega H(\boldsymbol{n}) + \sum_{x,y} \frac{\partial}{\partial n_{xy}} A_{xy} \right.$$

$$+ \frac{1}{2} \sum_i \sum_{\substack{x,y \\ w,z}} \left(\frac{\partial^2}{\partial n_{xy} \partial n_{wz}} B_{xy}^{(i)} B_{wz}^{(i)} \right.$$

$$\left. \left. + \frac{\partial^2}{\partial n_{xy} \partial n_{wz}} C_{xy}^{(i)} C_{wz}^{(i)} \right) \right] P(\underline{\lambda}, \tau). \tag{42}$$

In the actual calculation, instead of solving this equation directly, we solve the Ito-type Langevin equations with respect to the parameters of the Fokker-Planck equation which reproduce the distribution of $P(\underline{\lambda}, \tau)$:[18)]

$$d\Omega = -\Omega H(\boldsymbol{n}) d\tau, \tag{43}$$

$$d\boldsymbol{n} = -\boldsymbol{A} d\tau + \sum_i \boldsymbol{B}^{(i)} dW_i + \sum_i \boldsymbol{C}^{(i)} dW_i', \tag{44}$$

where dW and dW' are Wiener increments which satisfy $\langle dW_i \rangle = \langle dW'_i \rangle = \langle dW_i dW'_j \rangle = 0$ and $\langle dW_i dW_j \rangle = \langle dW'_i dW'_j \rangle = \delta_{ij} d\tau$.

Any expectation values of physical observables are evaluated by using the trace properties in Eqs.(32) and (33). Let $\hat{O}$ be a general observable consisting of the ladder operators, then the expectation value of $\hat{O}$ becomes

$$\langle \hat{O} \rangle = \frac{\text{Tr}[\hat{\rho}\hat{O}]}{\text{Tr}[\hat{\rho}]} = \frac{\int d\underline{\lambda} P(\underline{\lambda}, \tau) \text{Tr}\left[\Lambda(\underline{\lambda})\hat{O}\right]}{\int d\underline{\lambda} P(\underline{\lambda}, \tau) \text{Tr}\left[\Lambda(\underline{\lambda})\right]}$$

$$= \frac{\int d\underline{\lambda} P(\underline{\lambda}, \tau) \Omega O(\boldsymbol{n})}{\int d\underline{\lambda} P(\underline{\lambda}, \tau) \Omega}. \quad (45)$$

In the GBMC method, the integration with the weight $P(\underline{\lambda}, \tau)$ is achieved alternatively by summing up over all the walkers of the Langevin equations (43) and (44), i.e.,

$$\langle \hat{O} \rangle = \frac{\sum_i \Omega_i O(\boldsymbol{n}_i)}{\sum_i \Omega_i}. \quad (46)$$

We now regard Ω as the weight of the importance sampling in the Monte Carlo procedure. Note that from Eq. (43), the formal solution of the weight Ω becomes

$$\Omega(\tau) = \exp\left[-\int_0^\tau d\tau' H(\boldsymbol{n}(\tau'))\right]. \quad (47)$$

Since "Green's function" $\boldsymbol{n}(\tau)$ and $H(\boldsymbol{n})$ are always real, the weight Ω remains positive. Hence the negative sign problem does not appear.

3.2 Numerical integration

When integrating the Langevin equations, one has to be careful about the type of the SDEs. Since Eq. (44) is Ito-type SDE, the numerical integration must be done by Ito integration.[18] Here, we introduce two schemes of the numerical integration. The simplest one is Euler-Maruyama scheme:[20]

$$x_{i+1} = x_i + A(x_i)\Delta\tau + B(x_i)\Delta W_i. \quad (48)$$

This scheme is faster than any other scheme but is not stable in general. For a more stable integration, we use a semi-implicit iterative scheme[19] :

$$x_{i+1} = x_i + A(x_{i+1})\Delta\tau + B(x_i)\Delta W_i. \quad (49)$$

To solve the SDE, we make a first guess $\tilde{x}_{i+1}$ by Euler-Maruyama scheme (48). Then, $\tilde{x}_{i+1}$ is substituted into the drift term of (49) iteratively until a self-consistent solution is found. For the parameter values of the Hubbard model we have chosen the time step $\Delta\tau = 0.001$, then only a few iterations are needed because the initial guess from the Euler-Maruyama scheme is already close to the final solution.

3.3 Sampling method

For an efficient calculation, the importance sampling is needed. From Eq. (46), Ω can be regarded as a weight. Thus we can construct an importance sampling method with respect to Ω. Corney and Drummond use the branching method for importance sampling.[12,13,21] The branching method works by cloning the samples whose weights are large and by killing whose weights are small.

Assaad et al. also use a similar reconfiguration method but their method keeps total population constant.[16,22]

Here we propose another method which we call "successive Metropolis method" (see Fig. 1). In contrast
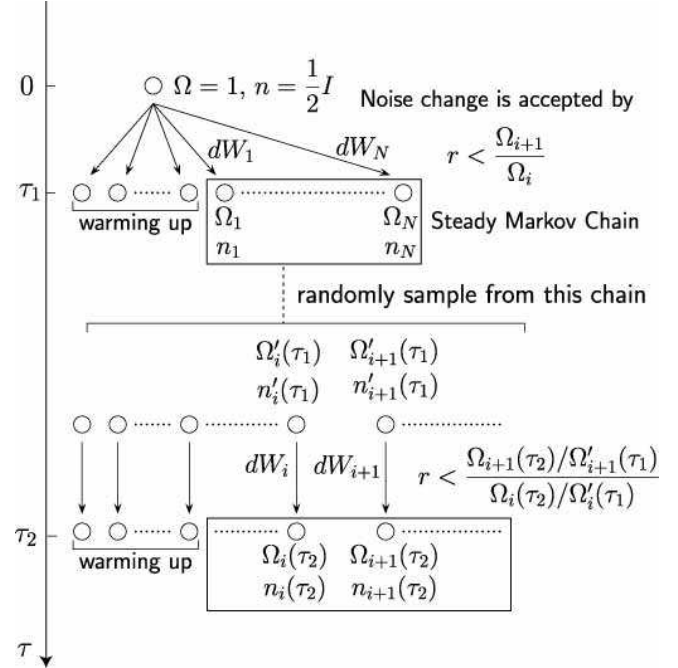

Fig. 1.　sampling method

with the usual Metropolis method, this method allows us to evolve τ successively. After a certain number of time steps $N\Delta\tau$, Monte Carlo samples are stored by the Metropolis algorithm with the following conditions.

· choose one starting point $(\Omega'_i(\tau_1), \boldsymbol{n}'_i(\tau_1))$

　and Wiener increments $d\vec{W}_i$

· evolve $N\Delta\tau$ steps to get $(\Omega_i(\tau_2), \boldsymbol{n}_i(\tau_2))$

· choose another starting point $(\Omega_{i+1}^{(\text{try})'}(\tau_1), \boldsymbol{n}_{i+1}^{(\text{try})'}(\tau_1))$

　and another Wiener increments $d\vec{W}_{i+1}^{(\text{try})}$

· evolve $N\Delta\tau$ steps to get $(\Omega_{i+1}^{(\text{try})}(\tau_2), \boldsymbol{n}_{i+1}^{(\text{try})}(\tau_2))$

· select the new sample with the condition that

$$(\Omega_{i+1}(\tau_2), \boldsymbol{n}_{i+1}(\tau_2), d\vec{W}_{i+1})$$

$$= \begin{cases} (\Omega_{i+1}^{(\text{try})}(\tau_2), \boldsymbol{n}_{i+1}^{(\text{try})}(\tau_2), d\vec{W}_{i+1}^{(\text{try})}) \\ \qquad\qquad \text{if } r < \frac{\Omega_{i+1}^{(\text{try})}(\tau_2)/\Omega_{i+1}^{(\text{try})'}(\tau_1)}{\Omega_i(\tau_2)/\Omega'_i(\tau_1)}, \\ (\Omega_i(\tau_2), \boldsymbol{n}_i(\tau_2), d\vec{W}_i) \qquad \text{otherwise,} \end{cases}$$

where r is a uniform random number distributed in $[0, 1)$. After a sufficient number of warming-up steps, the stored samples constitute a steady Markov chain which can be regarded as the new starting points of further time evolutions.

3.4 Systematic deviation

Here, we demonstrate some elementary results in the case of the two-site Hubbard model at $U/t = 4$ and $n = 1$ under the open boundary condition. Figure 2 shows the total energy, the specific heat C_v and the charge susceptibility $\chi_c = \frac{1}{NT}(\langle \hat{N}^2 \rangle - \langle \hat{N} \rangle^2)$, where $\hat{N} = \sum_{i\sigma} \hat{c}_{i\sigma}^\dagger \hat{c}_{i\sigma}$. All the numerical results show excellent agreement with the exact diagonalization result.
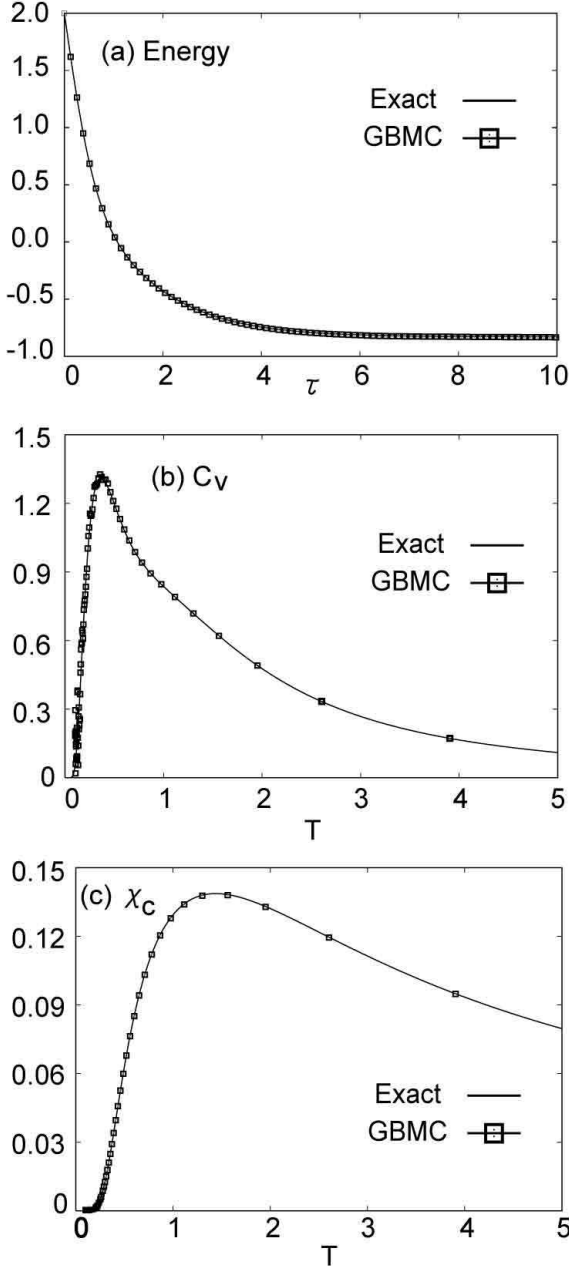
Fig. 2. (a) Energy of 2×1 lattice with $U/t = 4$ and $n = 1$ as a function of inverse temperature τ. Solid curve is obtained from the exact diagonalization. (b) Specific heat as a function of temperature T obtained from numerical derivative of (a). (c) Charge susceptibility as a function of temperature T.

However, simulation results deviate from the exact diagonalization results if the lattice size or the strength of the on-site interaction U becomes extremely larger.

Here, as an example we demonstrate the results for the case of the two-site Hubbard model at $U/t = 100$ and $n = 1$. Figure 3 shows the total energy and the to-
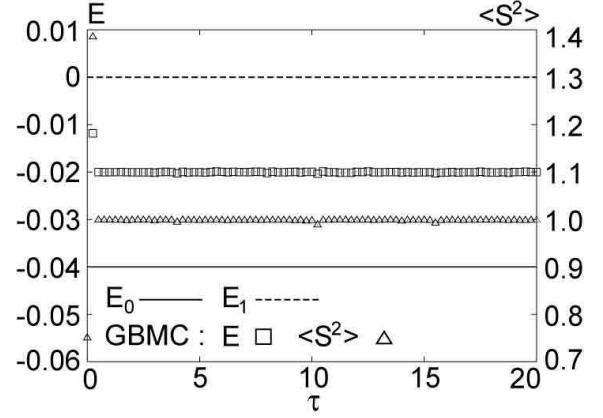

Fig. 3. Energy and total spin on 2×1 lattice at $U/t = 100$ and $n = 1$ as functions of inverse temperature τ. The squares and the triangles represent the GBMC results of the total energy and the total spin, respectively. Solid and dashed lines represent the exact value of the energy at the ground state E_0 and the first excited state E_1, respectively.

tal spin $\langle \hat{S}^2 \rangle$ for $U/t = 100$ and $n = 1$, where the total spin operator $\hat{S} = (\hat{S}^x, \hat{S}^y, \hat{S}^z)$ is defined as $\hat{S}^\alpha = \frac{1}{2} \sum_{i=1}^{N} \hat{c}_i^\dagger \sigma^\alpha \hat{c}_i$, $(\alpha = x, y, z)$. As is seen from Fig. 3, the energy obtained by the GBMC method is located just at the middle between the ground state and the triplet first excited states. This means that the GBMC method reproduces the state which is represented by the superposition of the ground state and the triplet first excited states. Indeed, the expectation value of the total spin is $\langle \hat{S}^2 \rangle = 1$, which is the middle point between the singlet state and the triplet states. Here, the ground state of the two-site Hubbard model at $n = 1$ is known to be represented as

$$|\Phi_0\rangle = \frac{1}{\sqrt{2}} \sqrt{1 + \frac{1}{\sqrt{1 + r^2}}} |\phi_{s1}\rangle$$

$$+ \frac{1}{2} \sqrt{1 - \frac{1}{\sqrt{1 + r^2}}} \left(|\phi_{s2}\rangle + |\phi_{s3}\rangle \right), \quad (50)$$

where $r = 4t/U$ and

$$|\phi_{s1}\rangle = \frac{1}{\sqrt{2}}(| \uparrow, \downarrow \rangle - | \downarrow, \uparrow \rangle) \qquad (51)$$

$$|\phi_{s2}\rangle = | \uparrow\downarrow, 0 \rangle, \quad |\phi_{s3}\rangle = |0, \uparrow\downarrow\rangle. \qquad (52)$$

Note that for $r \ll 1$, the ground state can be represented as

$$|\Phi_0\rangle \simeq |\phi_{s1}\rangle = \frac{1}{\sqrt{2}}(| \uparrow, \downarrow \rangle - | \downarrow, \uparrow \rangle). \qquad (53)$$

On the other hand, the triplet first excited states are represented as

$$|\phi_{t1}\rangle = \frac{1}{\sqrt{2}}(| \uparrow, \downarrow \rangle + | \downarrow, \uparrow \rangle) \qquad (54)$$

$$|\phi_{t2}\rangle = | \uparrow, \uparrow \rangle, \quad |\phi_{t3}\rangle = | \downarrow, \downarrow \rangle. \qquad (55)$$

The GBMC result of S^z is nearly zero in the whole range of τ (not shown), and the GBMC method converges with

$$|\phi_{1s}\rangle \pm |\phi_{1t}\rangle \propto |\uparrow, \downarrow\rangle \text{ or } |\downarrow, \uparrow\rangle. \qquad (56)$$

This means that samples obtained by the GBMC method are trapped in a symmetry broken states as $|\uparrow, \downarrow\rangle$ or $|\downarrow, \uparrow\rangle$, because the electron hopping is prohibited owing to the energy loss by the strong Coulomb repulsion.

When the system size becomes larger, the systematic deviation caused by the same reason as the case of the two-site Hubbard model at $U/t = 100$ occurs at a relatively small U/t. Here, for example we demonstrate the results for the case of the Hubbard model on the 4×1 ring with $n = 1$ under the periodic boundary condition in the x direction. Hereafter, the 4×1 lattice results are all obtained from the same boundary condition. Figure 4 shows the total energy in the case of $U/t = 1$ and $U/t = 4$. In the case of $U/t = 1$, the total energy agrees with the exact diagonalization result, while in the case of $U/t = 4$, the total energy deviates from the result of the exact diagonalization systematically. As is shown in the Table I, 4×1 lattice Hubbard model with $U/t = 4$ and $n = 1$ has triplet exicited states with the energy $E = -1.806424$. We conclude that the GBMC samples are trapped in a symmetry broken state constructed from a linear combination of the singlet ground state and the triplet excited states. To confirm this, we have calculated

Table I. Energy spectra of 4×1 lattice with $U/t = 4$ and $n = 1$ under the periodic boundary condition in the one-dimensional direction obtained by the exact diagonalization.

Energy	Number of Degeneracy
-2.102748	1
-1.806424	3
-1.068140	1
$\vdots$	$\vdots$

the expectation value of the total spin $\langle \hat{S}^2 \rangle$. As is shown in Fig. 5, the total spin has a nonzero value and these overlaps with excited $S > 0$ sectors cause the systematic deviation.

Fig. 5. Expectation value of total spin as a function of inverse temperature τ for the same case as Fig. 4 (b).

3.5 Power-law tails

To investigate the reason of the systematic deviation in detail, we calculate the distributions of the parameters of the Fokker-Planck equation (42) to analyze whether or not the boundary term in the partial integration of Eq. (35) exists.

First, we calculate the distribution of the weight Ω. As is seen form Fig. 6, the Monte Carlo step dependence of $P(\Omega)$ implies that the distributions of both $U/t = 1$ and $U/t = 4$ have upper limits of Ω, even though they have broad peak structures. To analyze this cutoff of $P(\Omega)$ in detail, we also calculate the integrated distribution defined by

$$Q(\Omega) = 1 - \int_0^\Omega P(\Omega')d\Omega'. \qquad (57)$$

As we see from Fig. 7, the integrated distributions of both $U/t = 1$ and $U/t = 4$ have cutoffs around $\Omega \simeq 10^{53}$ and $\Omega \simeq 10^{124}$, respectively. Thus there exists no boundary term with respect to the weight Ω.

Next, we calculate the distribution of Green's function $\boldsymbol{n}$. Figure 8 shows the distribution of Green's function. The abscissa N denotes the square root of the sum of
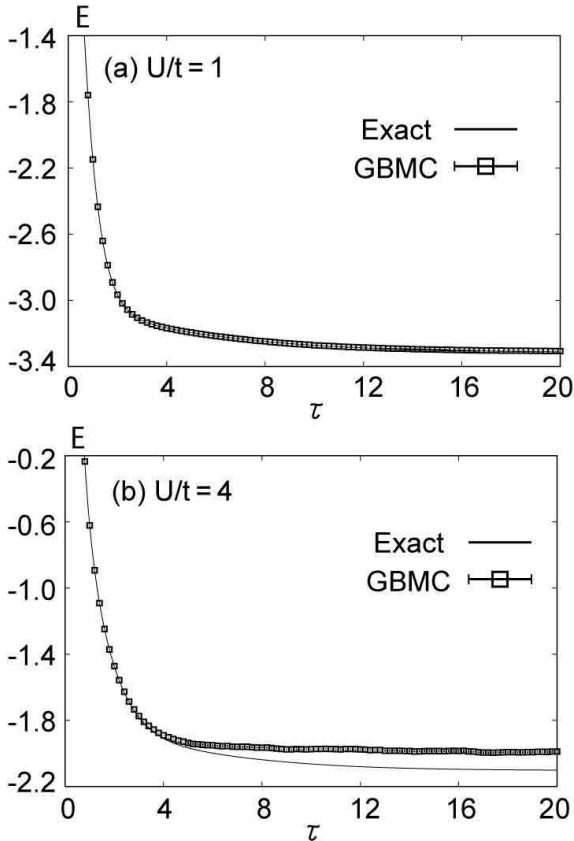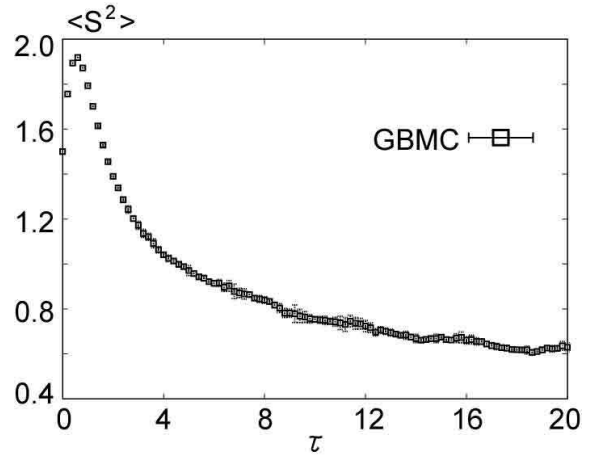

Fig. 4. (a) Energy of 4×1 lattice with $U/t = 1$ and $n = 1$ as a function of inverse temperature τ. Solid curve is obtained from the exact diagonalization. (b) Energy of 4×1 lattice with $U/t = 4$ and $n = 1$.
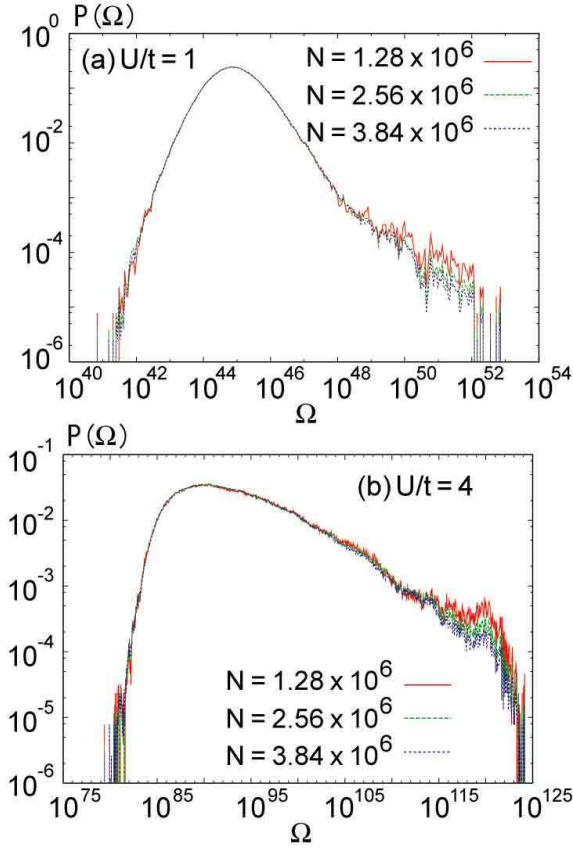
Fig. 6. (color online): Distribution of weight Ω at $\tau = 20$ for the 4×1 lattice at $n = 1$ in case of (a) $U/t = 1$ and (b) $U/t = 4$. Red (solid), green (dashed) and blue (dotted) curves represent the distribution obtained by 1.28×10^6, 2.56×10^6 and 3.84×10^6 Monte Carlo steps, respectively.
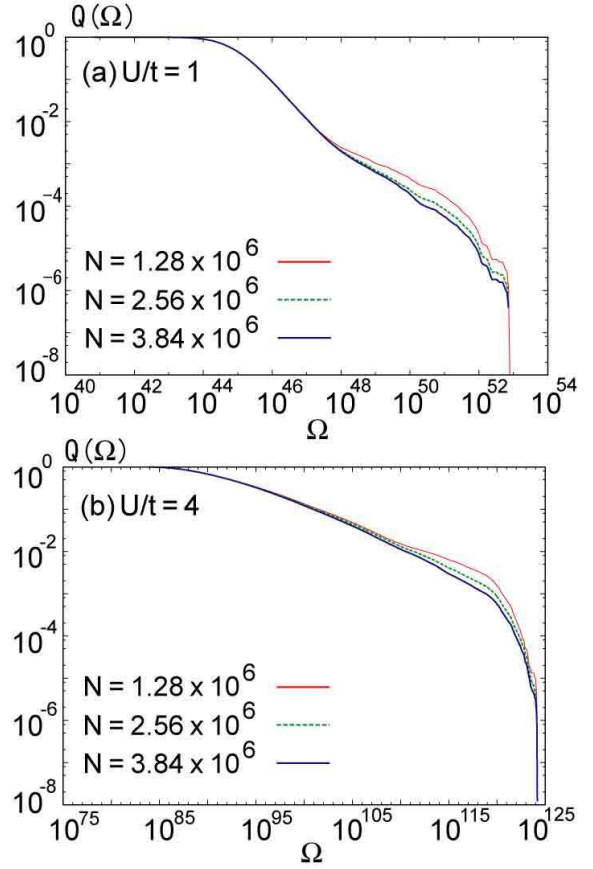

Fig. 7. (color online): Integrated distribution of weight Ω at $\tau = 20$ for the case of 4×1 lattice at $n = 1$ in case of (a) $U/t = 1$ and (b) $U/t = 4$. Red (light), green (dashed) and blue (dark) curves represent the distribution obtained by 1.28×10^6, 2.56×10^6 and 3.84×10^6 Monte Carlo steps, respectively.

each squared Green's function element, *i.e.*,

$$N = \sqrt{\sum_{ij\sigma} n^2_{(i\sigma),(j\sigma)}}, \qquad (58)$$

and the ordinate $P(N)$ denotes its probability distribution. Similarly to the case of $P(\Omega)$, the distributions $P(N)$ of both $U/t = 1$ and $U/t = 4$ have the upper limits. Thus there is no boundary term with respect to Green's function $\boldsymbol{n}$, either. However, as is seen from Fig. 8, the distribution tails of both $U/t = 1$ and $U/t = 4$ show power-law-like behaviors, *i.e.*, $P(N) \propto N^{-r}$, below their cutoffs. If the exponent of the power law r is less or equal to $p + 1$, the p-th moment of Green's function diverges, *i.e.*,

$$\int N^p P(N) dN \propto \int N^{p-r} dN \to \infty, \quad \text{if } r \leq p+1. \qquad (59)$$

To estimate the exponent of the power law in detail, we have also calculated the integrated distribution defined as

$$Q(N) = 1 - \int_0^N P(N') dN'. \qquad (60)$$

Figure 9 shows the integrated distribution of both $U/t = 1$ and $U/t = 4$. From the tails of $Q(N)$, we obtain the power-law exponent of $P(N)$, $r = 5.42 \pm 0.03$ for $U/t = 1$ and $r = 3.84 \pm 0.01$ for $U/t = 4$, respectively. To make the energy be well defined, the power-law exponent r must be larger than three, since the energy is the second-order moment of Green's function (see Eq. (59)). In the case of $U/t = 4$, the exponent r is larger than three, which also supports the absence of the boundary terms in the partial integration of Eq. (35). From the analysis of the distribution $P(\Omega)$ and $P(N)$, we conclude that the systematic deviations observed in the results of the original GBMC method are caused not from the boundary terms but from the trap to the quasi-stable states. In the next section, we introduce the quantum-number projection method which can remove the systematic deviation.

4. Quantum-Number Projection

Generally speaking, quantum many-body systems have several symmetries inherent in a Hamiltonian such as translational symmetry, U(1) symmetry, SU(2) symmetry, point group symmetry of the lattice, *etc.* Although these symmetries are sometimes broken in the thermodynamic limit, they must be preserved in finite size systems. In actual numerical calculations, however, these symmetries are not always preserved in restricted Hilbert space and the numerical calculation often suffers from system-
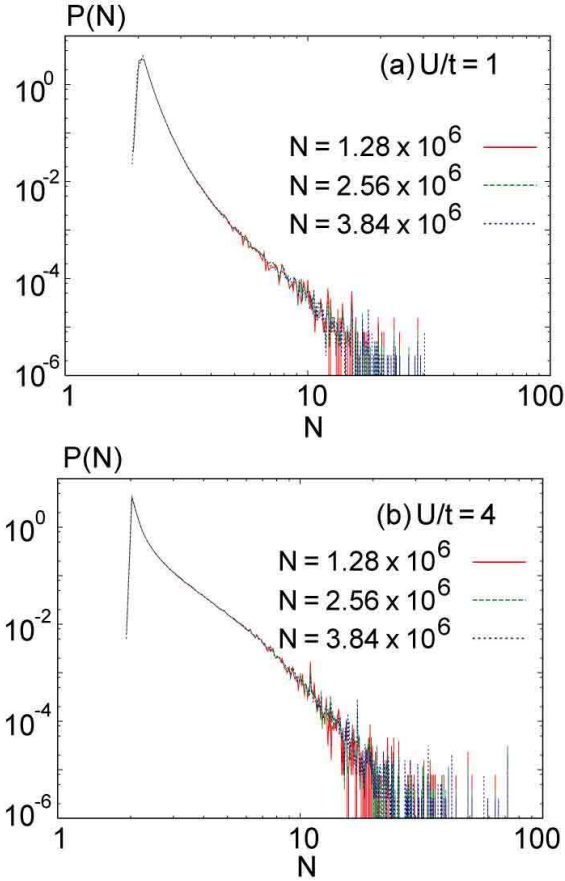
Fig. 8. (color online): Distribution of Green's function at $\tau = 20$ for the 4×1 lattice at $n = 1$ in case of (a) $U/t = 1$ and (b) $U/t = 4$. Red (solid), green (dashed) and blue (dotted) curves represent the distribution obtained by 1.28×10^6, 2.56×10^6 and 3.84×10^6 Monte Carlo steps, respectively.

atic errors.

One of the most promising device to restore these symmetries is the quantum-number projection which has been used successfully in the framework of the path-integral renormalization group method.[11] Also in the framework of the GBMC method, Assaad *et al.* have used the quantum-number projection.[16, 17] They proposed to project the density matrix onto given quantum numbers of the ground state after the ordinary GBMC sampling is performed and reproduced accurate ground states in some parameter regions. In this paper, we call their method post-projected GBMC (GBMC-PS) method.

In this section, we first review the mathematical framework of the quantum-number projection method, then introduce an alternative method for performing the quantum-number projection. In this method, we combine the quantum-number projection scheme in conjunction with the importance sampling of the original GBMC method. This allows us to perform the importance sampling with respect to the quantum-number-projected distribution, which makes it possible to treat wider parameter region than the previous studies.[16, 17] In this paper, we call this method pre-projected GBMC (PR-GBMC) method.
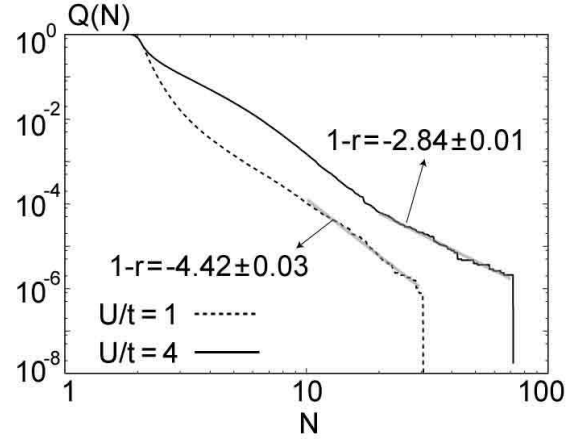

Fig. 9. Integrated distributions at $\tau = 20$ for the 4×1 lattice at $n = 1$. Dashed curve represents the distribution of $U/t = 1$, whereas solid curve is that of $U/t = 4$. The logarithmic fit to the tails leads to the exponent $r = 5.42 \pm 0.03$ for $U/t = 1$ and $r = 3.84 \pm 0.01$ for $U/t = 4$. Both data are obtained from 3.84×10^6 Monte Carlo steps.

4.1 Unitary transformation of a Gaussian operator

Before introducing the quantum-number projectors, we define a unitary transformation of a Gaussian operator which will be used in all the projectors. For any Hermitian matrix $\boldsymbol{h}(\boldsymbol{x})$, a Gaussian operator is transformed as[16, 17]

$$e^{i\hat{\boldsymbol{c}}^\dagger \boldsymbol{h}(\boldsymbol{x})\hat{\boldsymbol{c}}}\hat{\Lambda}(\Omega, \boldsymbol{n}) = \hat{\Lambda}(\Omega(\boldsymbol{x}), \boldsymbol{n}(\boldsymbol{x})), \quad (61)$$

where

$$\Omega(\boldsymbol{x}) = \Omega \det\left[(e^{i\boldsymbol{h}(\boldsymbol{x})} - \boldsymbol{I})\boldsymbol{n}^T + \boldsymbol{I}\right], \quad (62)$$

$$\boldsymbol{I} - \boldsymbol{n}(\boldsymbol{x}) = \left\{\left[(e^{i\boldsymbol{h}(\boldsymbol{x})} - \boldsymbol{I})\boldsymbol{n}^T + \boldsymbol{I}\right]^{-1}\right\}^T (\boldsymbol{I} - \boldsymbol{n}). \quad (63)$$

To prove Eqs.(62) and (63), we first introduce several identities of the Grassmann algebra[13] :

$$\langle \boldsymbol{\xi}|\boldsymbol{\xi}'\rangle = \exp\left[\overline{\boldsymbol{\xi}}\boldsymbol{\xi}' - \frac{1}{2}\overline{\boldsymbol{\xi}}\boldsymbol{\xi} - \frac{1}{2}\overline{\boldsymbol{\xi}'}\boldsymbol{\xi}'\right], \quad (64)$$

$$\langle \boldsymbol{\xi}| : A(\hat{\boldsymbol{c}}^\dagger, \hat{\boldsymbol{c}}) : |\boldsymbol{\xi}'\rangle = A(\overline{\boldsymbol{\xi}}, \boldsymbol{\xi}')$$
$$\times \exp\left[\overline{\boldsymbol{\xi}}\boldsymbol{\xi}' - \frac{1}{2}\overline{\boldsymbol{\xi}}\boldsymbol{\xi} - \frac{1}{2}\overline{\boldsymbol{\xi}'}\boldsymbol{\xi}'\right], \quad (65)$$

$$\boldsymbol{1} = \int \underbrace{\prod_x [d\overline{\xi}_x d\xi_x]}_{\mathcal{D}\boldsymbol{\xi}} |\boldsymbol{\xi}\rangle\langle\boldsymbol{\xi}|, \quad (66)$$

where $\boldsymbol{\xi}$ are Grassmann vectors and $|\boldsymbol{\xi}\rangle$ are Fermi coherent states.

In order to use the above identities, it is necessary to transform $e^{i\hat{\boldsymbol{c}}^\dagger \boldsymbol{h}\hat{\boldsymbol{c}}}$ into a normal ordered form. Because $\boldsymbol{h}$ is Hermitian, it can be diagonalized by the unitary transformation such that $\boldsymbol{h} = \boldsymbol{U}\boldsymbol{D}\boldsymbol{U}^\dagger$, where $\boldsymbol{U}$ is a unitary matrix and $\boldsymbol{D}$ is a diagonal one. With the canonical

transformation $\hat{\boldsymbol{\gamma}}^\dagger = \hat{\boldsymbol{c}}^\dagger \boldsymbol{U}$, $e^{i\hat{\boldsymbol{c}}^\dagger \boldsymbol{h}\hat{\boldsymbol{c}}}$ becomes

$$
\begin{aligned}
e^{i\hat{\boldsymbol{c}}^\dagger \boldsymbol{h}\hat{\boldsymbol{c}}} &= \prod_i e^{i\hat{\gamma}_i^\dagger \hat{\gamma}_i D_i} = \prod_i \left[1 + \sum_{n=1}^\infty \frac{(iD_i)^n}{n!} \hat{\gamma}_i^\dagger \hat{\gamma}_i \right] \\
&= \prod_i \left[1 + (e^{iD_i} - 1)\hat{\gamma}_i^\dagger \hat{\gamma}_i \right] \\
&= \prod_i \, : e^{(e^{iD_i}-1)\hat{\gamma}_i^\dagger \hat{\gamma}_i} :\; =: e^{\hat{\boldsymbol{c}}^\dagger (e^{i\boldsymbol{h}}-\boldsymbol{I})\hat{\boldsymbol{c}}} :, \quad (67)
\end{aligned}
$$

where $(\hat{\gamma}_i^\dagger \hat{\gamma}_i)^n = \hat{\gamma}_i^\dagger \hat{\gamma}_i$ is used. Then for any matrix $\boldsymbol{B}$, we obtain

$$
e^{i\hat{\boldsymbol{c}}^\dagger \boldsymbol{h}\hat{\boldsymbol{c}}} : e^{\hat{\boldsymbol{c}}^\dagger \boldsymbol{B}\hat{\boldsymbol{c}}} :
$$

$$
=: e^{\hat{\boldsymbol{c}}^\dagger (e^{i\boldsymbol{h}}-\boldsymbol{I})\hat{\boldsymbol{c}}} :: e^{\hat{\boldsymbol{c}}^\dagger \boldsymbol{B}\hat{\boldsymbol{c}}} :
$$

$$
= \int \mathcal{D}\boldsymbol{\xi}\mathcal{D}\boldsymbol{\eta}\mathcal{D}\boldsymbol{\gamma} |\boldsymbol{\xi}\rangle\langle\boldsymbol{\xi}| : e^{\hat{\boldsymbol{c}}^\dagger (e^{i\boldsymbol{h}}-\boldsymbol{I})\hat{\boldsymbol{c}}} : |\boldsymbol{\eta}\rangle\langle\boldsymbol{\eta}| : e^{\hat{\boldsymbol{c}}^\dagger \boldsymbol{B}\hat{\boldsymbol{c}}} : |\boldsymbol{\gamma}\rangle\langle\boldsymbol{\gamma}|
$$

$$
= \int \mathcal{D}\boldsymbol{\xi}\mathcal{D}\boldsymbol{\eta}\mathcal{D}\boldsymbol{\gamma} |\boldsymbol{\xi}\rangle
$$

$$
\times \exp[\overline{\boldsymbol{\xi}}(e^{i\boldsymbol{h}} - \boldsymbol{I})\boldsymbol{\eta}] \exp\left[\overline{\boldsymbol{\xi}}\boldsymbol{\eta} - \frac{1}{2}\overline{\boldsymbol{\xi}}\boldsymbol{\xi} - \frac{1}{2}\overline{\boldsymbol{\eta}}\boldsymbol{\eta}\right]
$$

$$
\times \exp[\overline{\boldsymbol{\eta}}\boldsymbol{B}\boldsymbol{\gamma}] \exp\left[\overline{\boldsymbol{\eta}}\boldsymbol{\gamma} - \frac{1}{2}\overline{\boldsymbol{\eta}}\boldsymbol{\eta} - \frac{1}{2}\overline{\boldsymbol{\gamma}}\boldsymbol{\gamma}\right] \langle\boldsymbol{\gamma}|
$$

$$
= \int \mathcal{D}\boldsymbol{\xi}\mathcal{D}\boldsymbol{\epsilon}\mathcal{D}\boldsymbol{\gamma} |\boldsymbol{\xi}\rangle \exp\left[\overline{\boldsymbol{\xi}}\boldsymbol{\epsilon} - \frac{1}{2}\overline{\boldsymbol{\xi}}\boldsymbol{\xi} - \frac{1}{2}\overline{\boldsymbol{\epsilon}}\boldsymbol{\epsilon}\right]
$$

$$
\times \exp\left[\overline{\boldsymbol{\epsilon}}\{e^{i\boldsymbol{h}}(\boldsymbol{B} + \boldsymbol{I}) - \boldsymbol{I}\}\boldsymbol{\gamma}\right]
$$

$$
\times \exp\left[\overline{\boldsymbol{\epsilon}}\boldsymbol{\gamma} - \frac{1}{2}\overline{\boldsymbol{\epsilon}}\boldsymbol{\epsilon} - \frac{1}{2}\overline{\boldsymbol{\gamma}}\boldsymbol{\gamma}\right] \langle\boldsymbol{\gamma}|
$$

$$
=: e^{\hat{\boldsymbol{c}}^\dagger [e^{i\boldsymbol{h}}(\boldsymbol{B}+\boldsymbol{I})-\boldsymbol{I}]\hat{\boldsymbol{c}}} :, \quad (68)
$$

where $\boldsymbol{\epsilon} = e^{i\boldsymbol{h}}\boldsymbol{\eta}$ and

$$
\mathcal{D}\boldsymbol{\epsilon} = \prod_x d\overline{\epsilon}_x d\epsilon_x = \prod_x d\overline{\eta}_x e^{-i\boldsymbol{h}} e^{i\boldsymbol{h}} d\eta_x = \mathcal{D}\boldsymbol{\eta}. \quad (69)
$$

Thus Eq. (61) is proven by taking $\boldsymbol{B} = -[2\boldsymbol{I} + (\boldsymbol{n}^T - \boldsymbol{I})^{-1}]$.

4.2 Examples of quantum-number projector
4.2.1 Particle-number projector

Since the GBMC method is a grand canonical approach, particle-number projection is needed when treating the canonical ensemble. The projector onto a state with a given particle-number N is defined as

$$
\begin{aligned}
\hat{P}_{\mathcal{N}}(N) &= \frac{1}{2\pi} \int_0^{2\pi} d\phi \langle N|\hat{T}_{\mathcal{N}}(\phi)|N\rangle^\dagger \hat{T}_{\mathcal{N}}(\phi) \\
&= \frac{1}{2\pi} \int_0^{2\pi} d\phi g_{\mathcal{N}}(\phi, N) e^{i\hat{\boldsymbol{c}}^\dagger \boldsymbol{h}_{\mathcal{N}}(\phi)\hat{\boldsymbol{c}}}, \quad (70)
\end{aligned}
$$

where $g_{\mathcal{N}}(\phi, N) = e^{-i\phi N}$ and

$$
\hat{T}_{\mathcal{N}}(\phi) = e^{i\phi \sum_i \hat{c}_i^\dagger \hat{c}_i}, \quad e^{i\boldsymbol{h}_{\mathcal{N}}(\phi)} = e^{i\phi}\boldsymbol{I}. \quad (71)
$$

Similarly, the projection onto the state which has N_σ electrons with spin σ is defined as

$$
\begin{aligned}
\hat{P}_{\mathcal{N}_\sigma}(N_\sigma) &= \frac{1}{2\pi} \int_0^{2\pi} d\phi \langle N_\sigma|\hat{T}_{\mathcal{N}_\sigma}(\phi)|N_\sigma\rangle^\dagger \hat{T}_{\mathcal{N}_\sigma}(\phi) \\
&= \frac{1}{2\pi} \int_0^{2\pi} d\phi g_{\mathcal{N}_\sigma}(\phi, N_\sigma) e^{i\hat{\boldsymbol{c}}^\dagger \boldsymbol{h}_{\mathcal{N}_\sigma}(\phi,\sigma)\hat{\boldsymbol{c}}}, \quad (72)
\end{aligned}
$$

where $g_{\mathcal{N}_\sigma}(\phi, N_\sigma) = e^{-i\phi N_\sigma}$ and

$$
\hat{T}_{\mathcal{N}_\sigma}(\phi) = e^{i\phi \sum_i \hat{c}_{i\sigma}^\dagger \hat{c}_{i\sigma}}, \quad e^{i\boldsymbol{h}_{\mathcal{N}_\sigma}(\phi,\sigma)} = e^{i\phi}\boldsymbol{I}_\sigma, \quad (73)
$$

$$
\boldsymbol{I}_\sigma = \delta_{\sigma,\uparrow} \begin{bmatrix} \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{bmatrix} + \delta_{\sigma,\downarrow} \begin{bmatrix} \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I} \end{bmatrix}. \quad (74)
$$

4.2.2 Total-spin projector

The SU(2) symmetry of the total spin is recovered by summing up all the Euler angles in the spin space.[11, 16, 17] Thus, projection onto a given total-spin state is defined as (here we restrict ourselves to the case of $S_z = 0$)

$$
\begin{aligned}
\hat{P}_{\mathcal{S}}(S) &= \frac{2S+1}{\int d\boldsymbol{\omega}} \int d\boldsymbol{\omega} \langle S, 0|\hat{T}_{\mathcal{S}}(\boldsymbol{\omega})|S, 0\rangle^\dagger \hat{T}_{\mathcal{S}}(\boldsymbol{\omega}) \\
&= \frac{2S+1}{8\pi^2} \int_0^{2\pi} d\alpha \int_0^\pi d\beta \sin\beta \int_0^{2\pi} d\gamma \\
&\quad \times P_S(\cos\beta) e^{i\hat{\boldsymbol{c}}^\dagger \boldsymbol{h}_{S_z}(\alpha)\hat{\boldsymbol{c}}} e^{i\hat{\boldsymbol{c}}^\dagger \boldsymbol{h}_{S_y}(\beta)\hat{\boldsymbol{c}}} e^{i\hat{\boldsymbol{c}}^\dagger \boldsymbol{h}_{S_z}(\gamma)\hat{\boldsymbol{c}}},
\end{aligned}
$$
$$(75)$$

where $P_S(\cos\beta)$ is the Sth Legendre polynomial and

$$
\hat{T}_{\mathcal{S}}(\boldsymbol{\omega}) = e^{i\alpha\hat{S}^z} e^{i\beta\hat{S}^y} e^{i\gamma\hat{S}^z}, \quad (76)
$$

$$
\boldsymbol{h}_{S_z}(\alpha) = \frac{\alpha}{2} \begin{bmatrix} \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{0} & -\boldsymbol{I} \end{bmatrix}, \quad \boldsymbol{h}_{S_y}(\beta) = \frac{\beta}{2} \begin{bmatrix} \boldsymbol{0} & -i\boldsymbol{I} \\ i\boldsymbol{I} & \boldsymbol{0} \end{bmatrix}. \quad (77)
$$

Here, $\hat{S}^\alpha, (\alpha = x, y, z)$ corresponds to the total α-component of spin :

$$
\hat{S}^\alpha = \frac{1}{2} \sum_i \hat{c}_i^\dagger \boldsymbol{\sigma}^\alpha \hat{c}_i. \quad (78)
$$

Since the total-spin projection involves triple integrals, the computational cost is rather high. However, if one takes $N_\uparrow$ projection and $N_\downarrow$ projection before the total-spin projection, the integrations about Euler angles α and γ can be done analytically. Thus, the total-spin projection is reduced to

$$
\hat{P}_{\mathcal{S}}(S) = \frac{2S+1}{2} \int_0^\pi d\beta g_{\mathcal{S}}(\beta, S) e^{i\hat{\boldsymbol{c}}^\dagger \boldsymbol{h}_{S_y}(\beta)\hat{\boldsymbol{c}}}, \quad (79)
$$

where $g_{\mathcal{S}}(\beta, S) = \sin\beta \, P_S(\cos\beta)$.

4.2.3 Lattice-symmetry projector

When the Hamiltonian is invariant under certain geometrical transformations, such geometrical symmetry is recovered by summing up all the transformations. For instance, when treating square lattice systems, they have C_{4v} lattice symmetry (see Fig. 10). By assuming that the square lattice lies in the x-y plane, the $\pi/2$ rotations
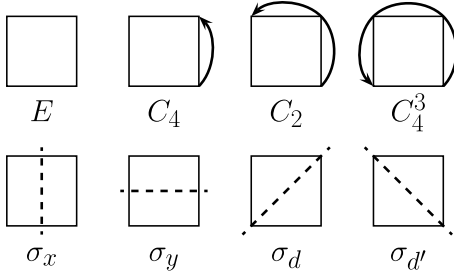
Fig. 10. Each element of C_{4v} group.

around the z-axis is achieved by the z-component of the angular momentum $\hat{L}_z$. Let $\hat{T}_{C_4} = e^{i\frac{\pi}{2}\hat{L}_z}$ such that

$$\hat{T}_{C_4} = e^{i\frac{\pi}{2}\hat{L}_z} = e^{i\hat{c}^\dagger \boldsymbol{h}_{C_4}\hat{c}} \tag{80}$$

$$\hat{T}_{C_4}\hat{c}_{i\sigma}^\dagger \hat{T}_{C_4}^{-1} = \hat{c}_{C_4(i)\sigma}^\dagger, \tag{81}$$

where $C_4(i)$ denotes a $\pi/2$ rotation around z-axis. From the above, $e^{i\hat{c}^\dagger \boldsymbol{h}_{C_4}\hat{c}}$ is represented as

$$\left[e^{i\hat{c}^\dagger \boldsymbol{h}_{C_4}\hat{c}}\right]_{(j\sigma'),(i\sigma)} = \delta_{\sigma\sigma'}\delta_{jC_4(i)}. \tag{82}$$

Similarly, the x-y mirror transformation $T_{\sigma_d} = e^{i\hat{c}^\dagger \boldsymbol{h}_{\sigma_d}\hat{c}}$ is defined as

$$\hat{T}_{\sigma_d}\hat{c}_{i\sigma}^\dagger \hat{T}_{\sigma_d}^{-1} = \hat{c}_{\sigma_d(i)\sigma}^\dagger, \tag{83}$$

where $\sigma_d(i)$ denotes a x-y mirror transformation of i-th site, $i.e.$, $\sigma_d(x,y) = (y,x)$. Then, the representation of σ_d becomes

$$\left[e^{i\hat{c}^\dagger \boldsymbol{h}_{\sigma_d}\hat{c}}\right]_{(j\sigma'),(i\sigma)} = \delta_{\sigma\sigma'}\delta_{j\sigma_d(i)}. \tag{84}$$

Although C_{4v} group has other elements as in Fig. 10, all can be generated by C_4 and σ_d (see Table.II). For example, $\sigma_x = C_4 \cdot \sigma_d$, $i.e.$,

$$C_4 \cdot \sigma_d \begin{pmatrix} x \\ y \end{pmatrix} = C_4 \begin{pmatrix} y \\ x \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} y \\ x \end{pmatrix}$$

$$= \begin{pmatrix} -x \\ y \end{pmatrix} = \sigma_x \begin{pmatrix} x \\ y \end{pmatrix} \tag{85}$$

Thus, the projection onto the C_{4v} symmetry sector reads

$$\hat{P}_{C_{4v}} = \frac{1}{8}\sum_\alpha g_{C_{4v}}(\alpha)\hat{T}_{C_{4v}}(\alpha), \tag{86}$$

where $\alpha = \{E, C_4, C_2, C_4^3, \sigma_x, \sigma_y, \sigma_d, \sigma_{d'}\}$ and

for s wave

$$g_{C_{4v}}(\alpha) = 1 \ (\forall\, \alpha), \tag{87}$$

for $d_{x^2-y^2}$ wave

$$g_{C_{4v}}(\alpha) = \begin{cases} 1 & (\alpha = E, C_2, \sigma_x, \sigma_y) \\ -1 & (\alpha = C_4, C_4^3, \sigma_d, \sigma_{d'}) \end{cases}, \tag{88}$$

for d_{xy} wave

$$g_{C_{4v}}(\alpha) = \begin{cases} 1 & (\alpha = E, C_2, \sigma_d, \sigma_{d'}) \\ -1 & (\alpha = C_4, C_4^3, \sigma_x, \sigma_y) \end{cases}. \tag{89}$$

For a simpler case, C_4-symmetry projector is obtained by restricting α to the subgroup of C_{4v}, $i.e.$,

$$\hat{P}_{C_4} = \frac{1}{4}\sum_\alpha g_{C_4}(\alpha)\hat{T}_{C_4}(\alpha), \tag{90}$$

$$\alpha = \{E, C_4, C_2, C_4^3\}, \tag{91}$$

where the definition of $g_{C_4}(\alpha)$ is same as $g_{C_{4v}}(\alpha)$ within $\alpha = \{E, C_4, C_2, C_4^3\}$.

Table II. The multiplication table of C_{4v}

	E	C_4	C_2	C_4^3	σ_x	σ_y	σ_d	$\sigma_{d'}$
E	E	C_4	C_2	C_4^3	σ_x	σ_y	σ_d	$\sigma_{d'}$
C_4	C_4	C_2	C_4^3	E	$\sigma_{d'}$	σ_d	σ_x	σ_y
C_2	C_2	C_4^3	E	C_4	σ_y	σ_x	$\sigma_{d'}$	σ_d
C_4^3	C_4^3	E	C_4	C_2	σ_d	$\sigma_{d'}$	σ_y	σ_x
σ_x	σ_x	σ_d	σ_y	$\sigma_{d'}$	E	C_2	C_4	C_4^3
σ_y	σ_y	$\sigma_{d'}$	σ_x	σ_d	C_2	E	C_4^3	C_4
σ_d	σ_d	σ_y	$\sigma_{d'}$	σ_x	C_4^3	C_4	E	C_2
$\sigma_{d'}$	$\sigma_{d'}$	σ_x	σ_d	σ_y	C_4	C_4^3	C_2	E

4.2.4 Total-momentum projector

When the Hamiltonian has the translational symmetry, the total momentum must be conserved. A translation by a lattice vector $\boldsymbol{R}$ is achieved by the operator $\hat{T}(\boldsymbol{R}) = e^{i\boldsymbol{R}\cdot \boldsymbol{k}_{\rm tot}}$. Here, $\boldsymbol{k}_{\rm tot}$ is defined by the Fourier transformation of the creation and the annihilation operators as

$$\boldsymbol{k}_{\rm tot} = \sum_{\boldsymbol{k},\sigma}\boldsymbol{k}\hat{c}_{\boldsymbol{k}\sigma}^\dagger \hat{c}_{\boldsymbol{k}\sigma}, \quad \hat{c}_{\boldsymbol{k}\sigma}^\dagger = \frac{1}{\sqrt{N}}\sum_i e^{i\boldsymbol{k}\cdot\boldsymbol{i}}\hat{c}_{i\sigma}^\dagger, \tag{92}$$

where $\boldsymbol{i}$ denotes a vector to the i-th site and N is the number of sites. The sum over $\boldsymbol{k}$ goes over all the points in the first Brillouin zone. The projection onto the Hilbert space with the total momentum $\boldsymbol{K}_0$ then reads:

$$\hat{P}_{\boldsymbol{\mathcal{K}}}(\boldsymbol{K}_0) = \frac{1}{N}\sum_{\boldsymbol{R}}\langle \boldsymbol{K}_0|\hat{T}_{\boldsymbol{\mathcal{K}}}(\boldsymbol{R})|\boldsymbol{K}_0\rangle^\dagger \hat{T}_{\boldsymbol{\mathcal{K}}}(\boldsymbol{R})$$

$$= \frac{1}{N}\sum_{\boldsymbol{R}}g_{\boldsymbol{\mathcal{K}}}(\boldsymbol{R},\boldsymbol{K}_0)e^{i\hat{c}^\dagger \boldsymbol{h}_{\boldsymbol{\mathcal{K}}}(\boldsymbol{R})\hat{c}}, \tag{93}$$

where $g_{\boldsymbol{\mathcal{K}}}(\boldsymbol{R},\boldsymbol{K}_0) = e^{-i\boldsymbol{R}\cdot\boldsymbol{K}_0}$ and the sum over $\boldsymbol{R}$ goes over all the lattice sites and $\boldsymbol{h}_{\boldsymbol{\mathcal{K}}}(\boldsymbol{R})$ is

$$(\boldsymbol{h}_{\boldsymbol{\mathcal{K}}}(\boldsymbol{R}))_{ij} = \frac{1}{N}\boldsymbol{R}\cdot\sum_{\boldsymbol{k}}\boldsymbol{k}e^{i\boldsymbol{k}\cdot(\boldsymbol{i}-\boldsymbol{j})}. \tag{94}$$

Then the matrix representation of $e^{i\hat{c}^\dagger \boldsymbol{h}_{\boldsymbol{\mathcal{K}}}(\boldsymbol{R})\hat{c}}$ reads

$$\left[e^{i\hat{c}^\dagger \boldsymbol{h}_{\boldsymbol{\mathcal{K}}}(\vec{R})\hat{c}}\right]_{(i\sigma),(j\sigma')} = \delta_{\sigma\sigma'}\langle\boldsymbol{i}|e^{i\boldsymbol{R}\cdot\boldsymbol{k}_{\rm tot}}|\boldsymbol{j}\rangle$$

$$= \delta_{\sigma\sigma'}\sum_{\boldsymbol{k}}\langle\boldsymbol{i}|\boldsymbol{k}\rangle e^{i\boldsymbol{R}\cdot\boldsymbol{k}}\langle\boldsymbol{k}|\boldsymbol{j}\rangle$$

$$= \delta_{\sigma\sigma'}\sum_{\boldsymbol{k}}e^{i\boldsymbol{k}\cdot(\boldsymbol{R}+\boldsymbol{i}-\boldsymbol{j})}$$

$$= \delta_{\sigma\sigma'}\delta_{\boldsymbol{i}+\boldsymbol{R},\boldsymbol{j}}. \tag{95}$$

4.3 Projected expectation values of observables

From the previous subsection, one can define a general quantum-number projector as

$$\hat{P} = \int d\boldsymbol{x} g(\boldsymbol{x}) \hat{T}(\boldsymbol{x}), \tag{96}$$

where $\hat{T}(\boldsymbol{x})$ is unitary and thus $\hat{P}^{\dagger} = \hat{P}$. For simplicity, we assume that the physical observable $\hat{O}$ commutes with $\hat{P}$, i.e., $\left[\hat{P}, \hat{O}\right]_{-} = 0$. Then, the projected expectation value of the observable $\hat{O}$ becomes

$$\langle \hat{O} \rangle_{\hat{P}} = \frac{\text{Tr}\left[\hat{P}\hat{\rho}\hat{P}\hat{O}\right]}{\text{Tr}\left[\hat{P}\hat{\rho}\hat{P}\right]} = \frac{\text{Tr}\left[\hat{P}\hat{\rho}\hat{O}\right]}{\text{Tr}\left[\hat{P}\hat{\rho}\right]}. \tag{97}$$

Here we use the projection property $\hat{P}^2 = \hat{P}$. Replacing the density-matrix operator by the sum over all the walkers $\rho = \int d\underline{\lambda} P(\underline{\lambda}) \hat{\Lambda}(\underline{\lambda}) \simeq \sum_i \hat{\Lambda}(\underline{\lambda}_i)$ yields

$$\langle \hat{O} \rangle_{\hat{P}} = \frac{\sum_i \int d\boldsymbol{x} \text{Tr}\left[\hat{T}(\boldsymbol{x})\hat{\Lambda}(\underline{\lambda}_i)\hat{O}\right]}{\sum_i \int d\boldsymbol{x} \text{Tr}\left[\hat{T}(\boldsymbol{x})\hat{\Lambda}(\underline{\lambda}_i)\right]}$$

$$= \frac{\sum_i \int d\boldsymbol{x} g(\boldsymbol{x})\Omega_i(\boldsymbol{x})O(\boldsymbol{n}_i(\boldsymbol{x}))}{\sum_i \int d\boldsymbol{x} g(\boldsymbol{x})\Omega_i(\boldsymbol{x})}, \tag{98}$$

where $\hat{T}(\boldsymbol{x})\hat{\Lambda}(\underline{\lambda}) = \hat{\Lambda}(\underline{\lambda}(\boldsymbol{x}))$.

4.4 Post-projected sampling method

In this section, we demonstrate the results of the GBMC-PS method. As we see in §3.4, the original GBMC method fails in reproducing the ground state of 4×1 lattice Hubbard model at $U/t = 4$ and $n = 1$. Here, we assess the accuracy obtained from the quantum-number projection to the density matrix obtained by the original GBMC method by following the idea of Assaad et al.[16, 17] Figure 11 shows the total energy compared with the exact diagonalization result. Here, we have projected
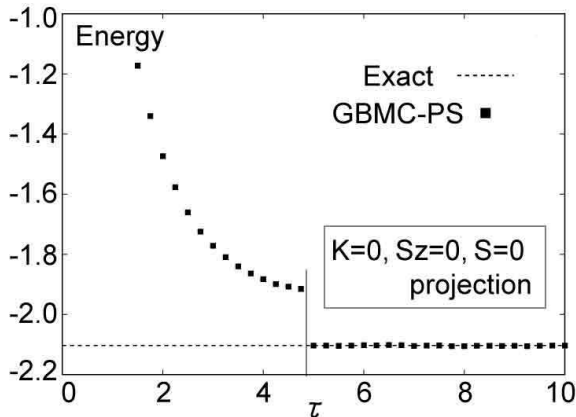

Fig. 11.   Energy of the 4×1-site Hubbard model with $U/t = 4$ and $n = 1$. Here, $\boldsymbol{K} = 0, S_z = 0$ and $S = 0$ projections are performed for $\tau > 5$. The integrals of S_z and S are evaluated by the Gauss-Legendre quadrature[23] with the summation over 10 points mesh each.

onto the state which has the total momentum $\boldsymbol{K} = 0$,

the total z-component of spin $S_z = 0$ and the total spin $S = 0$. In this case, the exact ground state energy is -2.102748 whereas our data is -2.1037 ± 0.0010. The error is obtained by averaging the data over the imaginary time after the convergence. This result shows that the quantum-number projection method well reproduces the ground state energy, which is not obtained in the framework of the original GBMC method. This result is consistent with the result of Assaad et al.[16, 17]

However, the GBMC-PS method suffers from a slow convergence when the interaction strength U/t becomes larger. As is illustrated in Fig. 12, the energy in the case of $U/t = 10$ and $n = 1$ on 4×4 lattice under the full periodic boundary condition obtained by the GBMC-PS method is not yet converged with the ground state at $\tau = 6$. Since the strong on-site repulsion U/t prevents
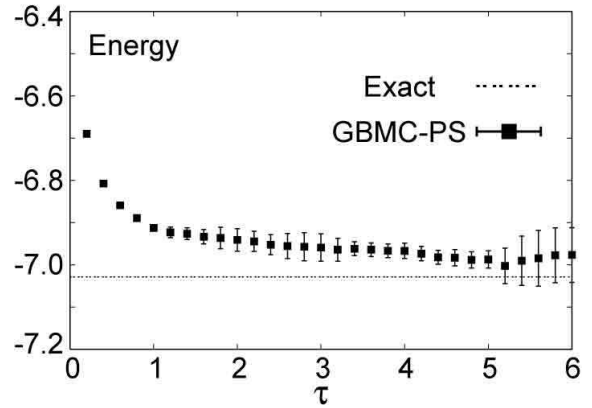

Fig. 12.   Energy as function of τ at $U/t = 10$ and $n = 1$ on the 4×4 lattice under the periodic boundary condition. Dashed line is the exact ground state energy obtained from the exact diagonalization. Here, for a fast convergence, we use Green's function obtained by the Hartree-Fock calculation as an alternative starting point. This scheme for the fast convergence will be discussed in §4.8 in detail.

the state from updating, the efficiency of the importance sampling of the original GBMC becomes worse with the increase of the interaction strength U/t. Thus, in the framework of the GBMC-PS method, it is difficult to store the samples which has a large overlap with the ground state. In the next subsection, we introduce the PR-GBMC method to overcome this slow convergence in the strong interaction regions.

4.5 Pre-projected sampling method

In this subsection, we introduce an alternative method for performing the quantum-number projection which is based on the importance sampling in combination with the quantum-number projection. This allows us to perform the sampling with the projected weight, which is more efficient than performing the sampling with the original weight. We call this pre-projection method the PR-GBMC method in contrast with the GBMC-PS method introduced by Assaad et al.[16, 17]

For the pre-projected sampling, we rewrite Eq. (98) as

$$\langle \hat{O} \rangle_{\hat{P}} = \sum_i \tilde{\Omega}_i \frac{\int d\boldsymbol{x} g(\boldsymbol{x}) \Omega_i(\boldsymbol{x}) O(\boldsymbol{n}_i(\boldsymbol{x}))}{\tilde{\Omega}_i} / \sum_i \tilde{\Omega}_i, \quad (99)$$

where $\tilde{\Omega}_i = \int d\boldsymbol{x} g(\boldsymbol{x}) \Omega_i(\boldsymbol{x})$. Estimating $\langle \hat{O} \rangle_{\hat{P}}$ is now reduced to the calculation of the weighted average of $\int d\boldsymbol{x} g(\boldsymbol{x}) \Omega_i(\boldsymbol{x}) O(\boldsymbol{n}_i(\boldsymbol{x}))/\tilde{\Omega}_i$ with respect to the projected weight $\tilde{\Omega}_i$.

The projected weight $\tilde{\Omega}_i$ stems from the projected density-matrix operator, i.e.,

$$\text{Tr}[\hat{P}\hat{\rho}\hat{P}^{\dagger}] = \sum_i \int d\boldsymbol{x} g(\boldsymbol{x}) \text{Tr}\left[\hat{T}(\boldsymbol{x}) \hat{\Lambda}(\Omega_i, \boldsymbol{n}_i) \right]$$

$$= \sum_i \int d\boldsymbol{x} g(\boldsymbol{x}) \Omega_i(\boldsymbol{x}) = \sum_i \tilde{\Omega}_i, \quad (100)$$

where $\hat{T}(\boldsymbol{x}) = e^{i\hat{c}^{\dagger} \boldsymbol{h}(\boldsymbol{x}) \hat{c}}$ and $\Omega_i(\boldsymbol{x}) = \Omega_i \det \left[(e^{i\boldsymbol{h}(\boldsymbol{x})} - \boldsymbol{I}) \boldsymbol{n}_i^T + \boldsymbol{I} \right]$. If the original samples have no overlap with the projected sector,

$$\text{Tr}[\hat{P}\hat{\rho}\hat{P}^{\dagger}] = \sum_i \tilde{\Omega}_i$$

$$= \sum_i \Omega_i \int d\boldsymbol{x} g(\boldsymbol{x}) \det \left[(e^{i\boldsymbol{h}(\boldsymbol{x})} - \boldsymbol{I}) \boldsymbol{n}_i^T + \boldsymbol{I} \right]$$

$$(101)$$

becomes zero. However, from Eq. (47) the unprojected weight Ω_i is always positive. It suggests that the factor $\int d\boldsymbol{x} g(\boldsymbol{x}) \det \left[(e^{i\boldsymbol{h}(\boldsymbol{x})} - \boldsymbol{I}) \boldsymbol{n}_i^T + \boldsymbol{I} \right]$ in Eq. (101) causes the reduction of Ω_i to $\tilde{\Omega}_i$ when the original samples have small overlap with the projected sector. Empirically, we find that this reduction is realized by the cancellation of positive $\tilde{\Omega}_i$ and negative $\tilde{\Omega}_i$, thus $\tilde{\Omega}_i$ is not positive definite. This is the source of the negative sign problem in the PR-GBMC method. In this case, we introduce the sign variable $S_i = \pm 1$ by $\tilde{\Omega}_i = S_i |\tilde{\Omega}_i|$, and the importance sampling is performed with the absolute value of $\tilde{\Omega}_i$, namely we calculate

$$\langle \hat{O} \rangle_{\hat{P}} = \sum_i |\tilde{\Omega}_i| S_i \frac{\int d\boldsymbol{x} g(\boldsymbol{x}) \Omega_i(\boldsymbol{x}) O(\boldsymbol{n}_i(\boldsymbol{x}))}{\tilde{\Omega}_i} / \sum_i |\tilde{\Omega}_i| S_i. \quad (102)$$

Appearance of the negative sign is evaluated by calculating the expectation value of the sign defined as

$$\langle S \rangle = \frac{\sum_i |\tilde{\Omega}_i| S_i}{\sum_i |\tilde{\Omega}_i|}. \quad (103)$$

Below we show results of the PR-GBMC method with the quantum-number projection $\boldsymbol{K} = 0$, $S_z = 0$ and $S = 0$. Figure 13 shows the energy of 4×1 lattice Hubbard model with $U/t = 4$ and $n = 1$ calculated by the PR-GBMC method. Averaging the data over imaginary time gives the energy -2.1026 ± 0.0002 in agreement with the exact diagonalization result -2.102748. During the simulation, the average sign $\langle S \rangle$ is kept unity, i.e., there is no negative sign. We will discuss later the negative sign problem in more detail.

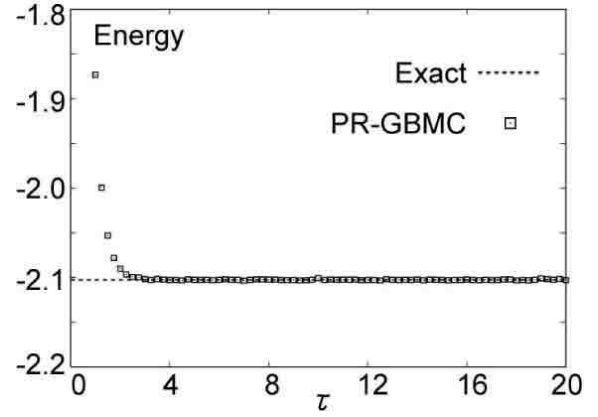One of the main advantages in the PR-GBMC method is that this method allows us to analyze directly the

Fig. 13. Energy of the 4×1-site lattice Hubbard model with $U/t = 4$ and $n = 1$. Here, $\boldsymbol{K} = 0$, $S_z = 0$ and $S = 0$ projections are performed.

change in the probability distributions caused by the quantum-number projection:

$$\frac{\partial P(\Omega, \boldsymbol{n})}{\partial \tau} = \mathcal{L} P(\Omega, \boldsymbol{n})$$

$$\xrightarrow{\text{projection}} \frac{\partial \tilde{P}(\tilde{\Omega}, \tilde{\boldsymbol{n}})}{\partial \tau} = \tilde{\mathcal{L}} \tilde{P}(\tilde{\Omega}, \tilde{\boldsymbol{n}}), \quad (104)$$

where the projected variables are denoted by tilde. The reweighted distribution $\tilde{P}$ cannot be calculated in the framework of the GBMC-PS method, because in the GBMC-PS method, the quantum-number projection is performed by reweighting the importance sampling of the original GBMC method, i.e.,

$$\langle \hat{O} \rangle_{\hat{P}} = \frac{\sum_i \int d\boldsymbol{x} g(\boldsymbol{x}) \Omega_i(\boldsymbol{x}) O(\boldsymbol{n}_i(\boldsymbol{x}))}{\sum_i \int d\boldsymbol{x} g(\boldsymbol{x}) \Omega_i(\boldsymbol{x})}$$

$$= \frac{\sum_i \Omega_i \frac{\int d\boldsymbol{x} g(\boldsymbol{x}) \Omega_i(\boldsymbol{x}) O(\boldsymbol{n}_i(\boldsymbol{x}))}{\Omega_i}}{\sum_i \Omega_i \frac{\int d\boldsymbol{x} g(\boldsymbol{x}) \Omega_i(\boldsymbol{x})}{\Omega_i}}$$

$$\xrightarrow{\text{sampling by } \Omega_i} \frac{\frac{1}{N_{\text{mcs}}} \sum_i \frac{\int d\boldsymbol{x} g(\boldsymbol{x}) \Omega_i(\boldsymbol{x}) O(\boldsymbol{n}_i(\boldsymbol{x}))}{\Omega_i}}{\frac{1}{N_{\text{mcs}}} \sum_i \frac{\int d\boldsymbol{x} g(\boldsymbol{x}) \Omega_i(\boldsymbol{x})}{\Omega_i}}, \quad (105)$$

where N_{mcs} is the number of Monte Carlo samples. As is seen from Eq. (105), what one can obtain by the GBMC-PS method is the distribution of $\int d\boldsymbol{x} g(\boldsymbol{x}) \Omega_i(\boldsymbol{x}) O(\boldsymbol{n}_i(\boldsymbol{x}))/\Omega_i$ and $\int d\boldsymbol{x} g(\boldsymbol{x}) \Omega_i(\boldsymbol{x})/\Omega_i$. Thus, the distribution of $\langle \hat{O} \rangle_{\hat{P}}$ itself can not be calculated by GBMC-PS method.

We now calculate the projected probability distribution $\tilde{P}(\tilde{N})$ to compare with the unprojected distribution $P(N)$ obtained by the original GBMC method. Figure 14 shows the distribution of Green's function. The abscissa N represents $\sqrt{\sum_{ij\sigma} n_{(i\sigma),(j\sigma)}^2}$ for the GBMC method and $\sqrt{\sum_{ij\sigma\sigma'} \tilde{n}_{(i\sigma),(j\sigma')}^2}$ for the PR-GBMC method, where

$$\tilde{\boldsymbol{n}} = \int d\boldsymbol{x} g(\boldsymbol{x}) \Omega(\boldsymbol{x}) \boldsymbol{n}(\boldsymbol{x})/|\tilde{\Omega}|. \quad (106)$$

As is illustrated in Fig. 14, the projected distribution $\tilde{P}(\tilde{N})$ decays exponentially. This means that the
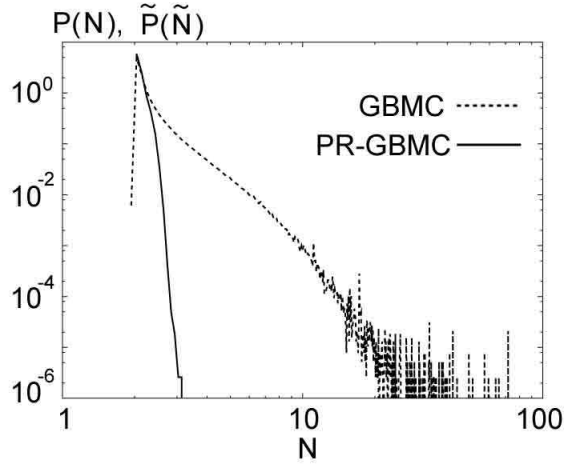
Fig. 14. Distribution of Green's function. The solid curve shows the distribution obtained by the PR-GBMC and the dashed curve shows the distribution obtained by the GBMC which is the same data as in Fig. 8(b). The data are obtained for the 4×1-site Hubbard model with $U/t = 4$ and $n = 1$ at $\tau = 20$ with 3.84×10^6 Monte Carlo steps.

quantum-number projection actually reduces the phase-space. In the PR-GBMC method, the importance sampling is performed with respect to this reduced phase-space. Therefore, the convergence to the ground state is faster than the GBMC-PS method which is based on the importance sampling with the original weight.

4.6 Comparison between GBMC-PS and PR-GBMC

In the previous subsection, we have confirmed that the quantum-number projection changes the probability distribution and removes errors arising in the original GBMC procedure. Here, we make a comparison between the GBMC-PS and PR-GBMC methods to discuss their merits and demerits.

Figure 15 shows the energy of the 4×4-site Hubbard model with $U/t = 4$ and $n = 1$ under the periodic boundary condition. As is seen from Fig. 15, the energy obtained by the PR-GBMC method converges with the ground state faster than that obtained by the GBMC-PS method. This comes from the difference of the sampling procedure. In GBMC-PS method importance sampling is performed with respect to the unprojected weight $\Omega = e^{-\int_0^\tau H(\boldsymbol{n})d\tau'}$. Thus the sampling depends only on the energy. On the other hand, PR-GBMC method makes use of the projected weight $\tilde{\Omega} = \Omega \int d\boldsymbol{x} g(\boldsymbol{x}) \det[(e^{i\boldsymbol{h}} - \boldsymbol{I})\boldsymbol{n}^T + \boldsymbol{I}]$, which reflects not only the energy but also the overlap with the projected sector. Therefore, the PR-GBMC method allows the convergence to the ground state at smaller τ than the GBMC-PS method. However, in the PR-GBMC method we have to perform the projection for every sample, while the GBMC-PS method requires the projection only for accepted samples, which makes the computation time shorter. Empirically, we find that the energy obtained by both the GBMC-PS and the PR-GBMC methods converges with the ground state when the on-site interaction U/t is not too large, whereas the PR-GBMC method

is more efficient. Actually the PR-GBMC method offers a better convergence at larger U/t. This is because the original GBMC sampling fails in making samples which have enough overlap with the ground state at relatively large U/t. This possibly causes a serious minus sign problem. Thus when treating large U/t systems (typically $U/t \gtrsim 4$), the PR-GBMC method has to be employed.
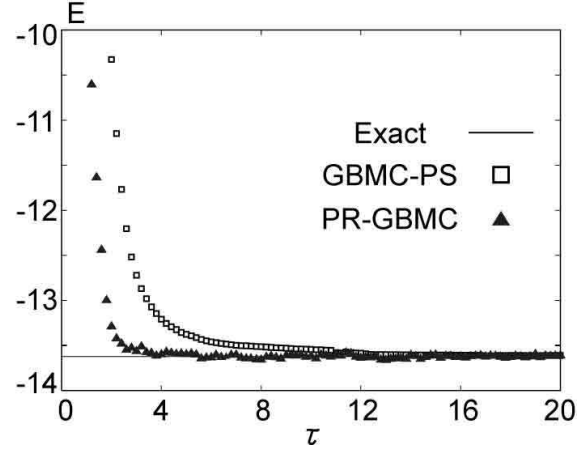

Fig. 15. Energy of the 4×4 lattice with $U/t = 4$ and $n = 1$ as functions of τ. The squares represent the result of the GBMC-PS method and the triangles represent the result of the PR-GBMC method.

4.7 Negative sign problem

Since the PR-GBMC method is based on samplings with respect to the projected weights $\tilde{\Omega}$, there appears minus sign if a sample has small overlap with the projected sector. Figure 16 shows PR-GBMC results of the total energy and the expectation value of the sign $\langle S \rangle = \sum_i |\tilde{\Omega}_i| S_i / \sum_i |\tilde{\Omega}_i|$ on the 4×4 lattice at $U/t = 4$ and $n = 1$. As we see in Fig. 16, the average sign decreases first, but it recovers along with the convergence of the energy. This means that if the overlap with the projected sector is small, the expectation value of the sign becomes small, while $\langle S \rangle$ recovers when the samples gain a large overlap with the quantum-number-projected state. We note that the τ dependence of $\langle S \rangle$ is completely different from that in the conventional AFQMC method as well as from that in other methods, where $\langle S \rangle$ exponentially decreases to zero with increasing τ.

4.8 Acceleration of convergence

In the previous subsections, we have proposed and studied the PR-GBMC method and examined the negative sign problem. In the PR-GBMC method, the negative sign appears when the overlap with the projected sector is small. To avoid this problem, instead of employing $\boldsymbol{n}_0 = \frac{1}{2}\boldsymbol{I}$ as the initial condition, it is better to use Green's function obtained by the Hartree-Fock calculation as an alternative starting point. Since the initial state is already that of the Hartree-Fock solution, the overlap with the ground state is expected to be relatively
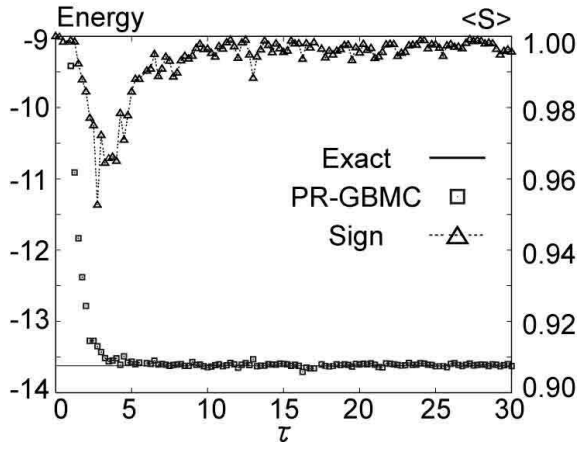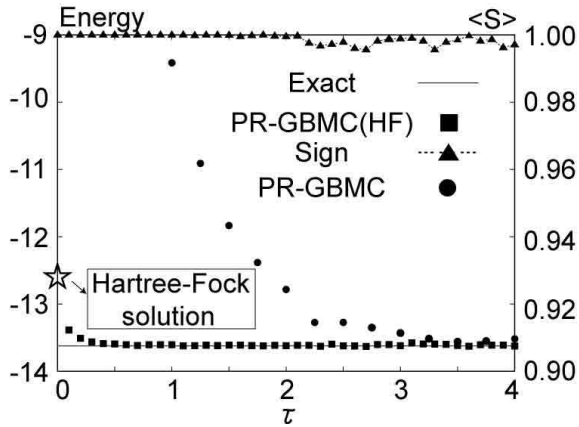
Fig. 16. Energy and average sign of PR-GBMC results as functions of τ. The squares represent the energy of 4×4 lattices under the periodic boundary condition with $U/t = 4$ and $n = 1$ and the triangles represent the average sign. The initial condition at $\tau = 0$ is Green's functions for the infinite temperature.

larger than $\boldsymbol{n}_0 = \frac{1}{2}\boldsymbol{I}$. Therefore, the convergence to the ground state becomes faster and the expectation value of the average sign becomes stable. Figure 17 shows the total energy of 4×4 lattices under the periodic boundary condition with $U/t = 4$ and $n = 1$ together with its average sign. As is seen from Fig. 17, the total energy converges already at $\tau = 0.5$ and the average sign is nearly unity in the whole range of τ.

Fig. 17. Energy and average sign of PR-GBMC results as functions of τ. The squares represent the energy of 4×4 lattices under the periodic boundary condition with $U/t = 4$ and $n = 1$ started from the Hartree-Fock solution (open star) and the triangles represent the average sign. For comparison, the PR-GBMC result obtained from the infinite temperature at $\tau = 0$ is plotted (circles).

4.9 Applicability of PR-GBMC method
4.9.1 U/t dependence of convergence

In the PR-GBMC method, the distributions of the phase-space variables $P(\Omega)$ and $P(N)$ are transformed to the projected distributions $\tilde{P}(\tilde{\Omega})$ and $\tilde{P}(\tilde{N})$ which decay faster than the original distributions. However, the decay

of the projected distributions becomes slower with the increase of U/t. In this subsection, by comparing the U/t dependence of the energy convergence with that of the projected distribution, we discuss the applicable range of the PR-GBMC method.

Figure 18 shows the energy of 4×4 lattice under the periodic boundary condition with $n = 1$ at $U/t = 4$, 10 and 15. All the results are obtained by 1.6×10^6 Monte Carlo steps with the pre-projection at $\boldsymbol{K} = 0$, $S_z = 0$ and $S = 0$. We employ the data at $\tau = 0$ from the solutions of the Hartree-Fock calculation with $U_{\mathrm{int}} = U/2$. Measurements are divided into 5 bins and the error bars are estimated by the variance among the 5 bins. As is
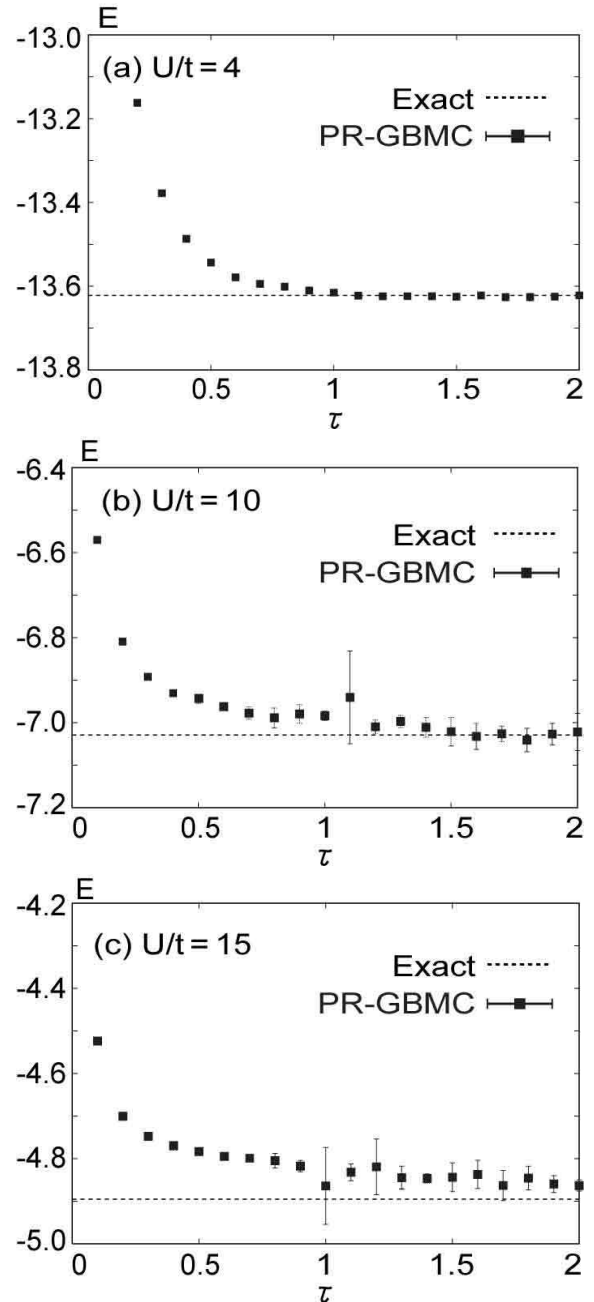

Fig. 18. Energy as functions of τ at (a) $U/t = 4$, (b) $U/t = 10$ and (c) $U/t = 15$ for the 4×4-site Hubbard model under the periodic boundary condition . Dashed line is the exact ground state energy obtained from the exact diagonalization.

seen from Fig. 18, the convergence with the ground state becomes slower with the increase of U/t and at $U/t = 15$, the energy obtained by the PR-GBMC method does not yet converge with the ground state at $\tau = 2$ because of the slow convergence. At $U/t = 15$, since the energy seems to show further decrease beyond $\tau = 2$, further time evolution is required whereas the statistical error becomes large. This large statistical error comes from the fast diffusion of the probability distribution of the sampling weight Ω. Since the definition of the weight is given by $\Omega = e^{-\int_0^\tau H(\boldsymbol{n})d\tau'}$, the weight Ω grows exponentially with τ. Thus, the phase-space to be sampled becomes larger if the convergence with the ground state becomes slower because of the strong interaction. This requires much more computation time, which determines the practical limitation of the PR-GBMC method.

To confirm this, we calculate the distributions of the projected phase-space variables $\tilde{\Omega}$ and $\tilde{N} = \sqrt{\sum_{ij\sigma\sigma'} \tilde{n}_{(i\sigma),(j\sigma')}}$. Figures 19 and 20 show the integrated distributions $\tilde{Q}(\tilde{\Omega})$ and $\tilde{Q}(\tilde{N})$ defined by

$$\tilde{Q}(\tilde{\Omega}) = 1 - \int_0^{\tilde{\Omega}} \tilde{P}(\tilde{\Omega}')d\tilde{\Omega}', \qquad (107)$$

$$\tilde{Q}(\tilde{N}) = 1 - \int_0^{\tilde{N}} \tilde{P}(\tilde{N}')d\tilde{N}'. \qquad (108)$$

As we see in Fig. 19, the Monte Carlo step dependence of $\tilde{Q}(\tilde{\Omega})$ implies the existence of the cutoffs in the distributions of $\tilde{\Omega}$ for all U/t. Thus, the boundary terms with respect to the projected weight $\tilde{\Omega}$ do not appear to exist for all U/t. However, the phase-space of the projected weight $\tilde{\Omega}$ becomes larger with U/t and at $U/t = 15$, and a distinct plateau structure with steps caused by the lack of the large $\tilde{\Omega}$ samples is seen in the tail of the distribution (see Fig. 19.c). This means that in the case of $U/t = 15$, 1.6×10^6 Monte Carlo steps are not enough for the accurate sampling of events at large weight Ω. Since the samples with large weight seldom appear, but contribute to lowering the energy, these samples cause the spike structure in the distribution of the energy (see Fig. 21). Thus, the statistical error of the energy becomes larger.

The slow convergence of the distribution due to the strong interaction is also observed in the distribution of the projected Green's function $\tilde{Q}(\tilde{N})$. As is seen from Fig. 20, the convergence of the distribution tail becomes slower as U/t increases. Although the energy converges with the ground state at $U/t = 10$ as in Fig. 18.b, the plateau structure in $\tilde{Q}(\tilde{N})$ visible at $U/t = 10$ (Fig. 20.b) signals the slow convergence in the PR-GBMC method. However, the convergence of the energy at $U/t = 10$ shows that a small plateau structure arising in the tail part of $\tilde{Q}(\tilde{N})$ does not yet cause a bad effect on the convergence of the energy. At $U/t = 15$, the distribution tail of $\tilde{Q}(\tilde{N})$ shows no indication of the existence of the cutoff at least up to 1.6×10^6 Monte Carlo steps, which indicates that the number of Monte Carlo steps 1.6×10^6 is not enough at $U/t = 15$.

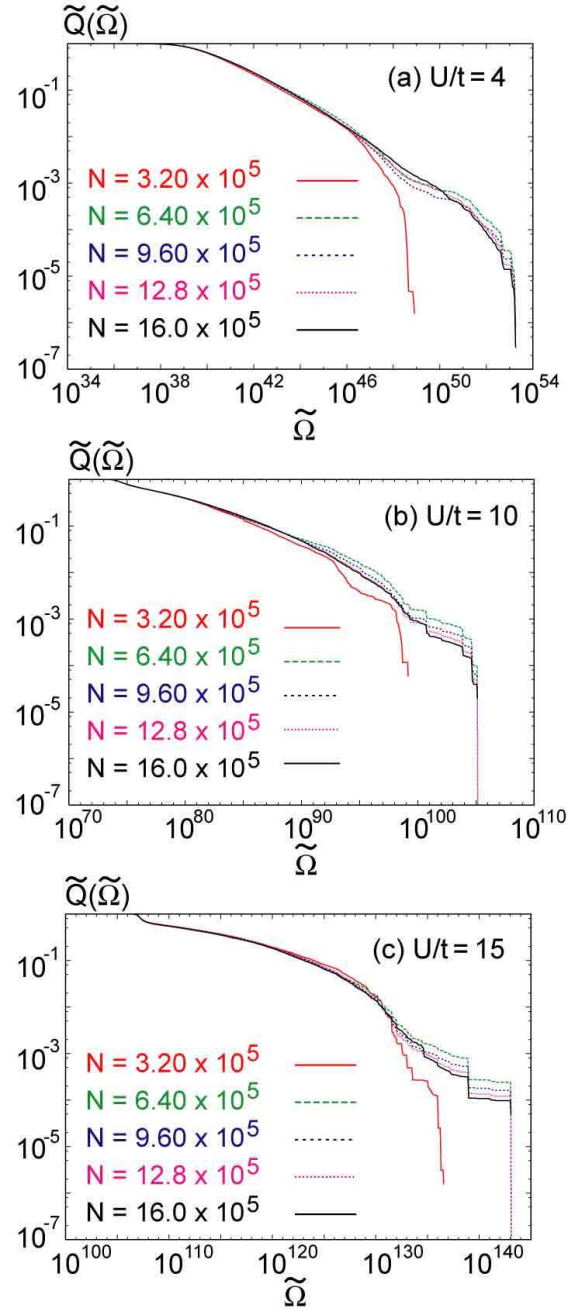These slow convergences of the distributions at large U/t come from the expansion of the phase-space to be

Fig. 19. (color online) Integrated distribution of projected weight $\tilde{\Omega}$ at $\tau = 2$ for the 4×4-site square lattice under the periodic boundary condition with (a) $U/t = 4$, (b) $U/t = 10$ and (c) $U/t = 15$. In all the panels, red (light), green (long dashed), blue (dashed), pink (dotted) and black (dark) curves represent the distribution obtained by 3.20×10^5, 6.40×10^5, 9.60×10^5, 12.8×10^5 and 16.0×10^5 Monte Carlo steps, respectively.

sampled. From Eqs.(38-40), the drift term and the diffusion terms in Langevin equation (44) are proportional to U and $\sqrt{U}$, respectively. Thus the diffusion speed of the distributions becomes faster with the increase of U/t, which means that the phase-space to be sampled becomes larger. This larger sampling space requires larger Monte Carlo steps and causes an insufficient sampling at large U/t. Therefore, it is advisable to monitor the convergence of the distributions $\tilde{Q}(\tilde{\Omega})$ and $\tilde{Q}(\tilde{N})$ to ensure that the number of Monte Carlo steps is sufficiently
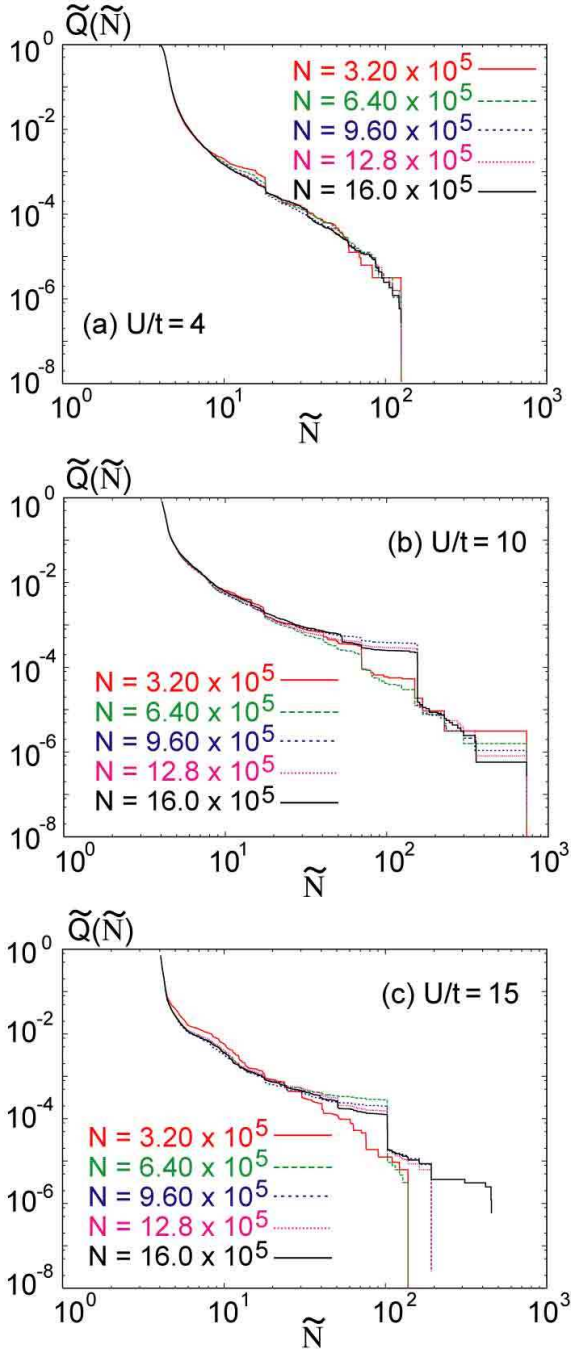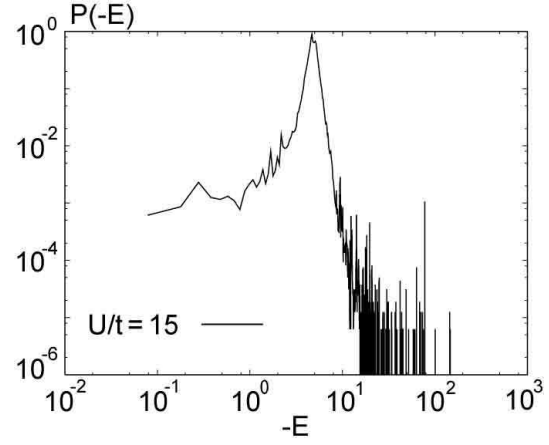
Fig. 21. Distribution of energy at $U/t = 15$ for 4×4-site Hubbard model under the periodic boundary condition at $n = 1$. The data are obtained from 1.6×10^6 Monte Carlo steps at $\tau = 2$

Fig. 20. (color online): Integrated distribution of projected Green's function $\tilde{N}$ at $\tau = 2$ for the 4×4-site square lattice under the periodic boundary condition with (a) $U/t = 4$, (b) $U/t = 10$ and (c) $U/t = 15$. In all the panels, red (light), green (long dashed), blue (dashed), pink (dotted) and black (dark) curves represent the distributions obtained from 3.20×10^5, 6.40×10^5, 9.60×10^5, 12.8×10^5 and 16.0×10^5 Monte Carlo steps, respectively.

large.

4.9.2 Size dependence of convergence

In this subsection, we analyze the size dependence of the convergence by employing the PR-GBMC method with $\boldsymbol{K} = 0$, $S_z = 0$ and $S = 0$ projections. Figure 22 shows the energy of 6×6, 8×8 and 10×10 lattices at $U/t = 4$ and $n = 1$. As is seen from Fig. 22,

all the simulation results converge with the ground state energy obtained by the AFQMC method with the Trotter discretization of $\Delta\tau = 0.025$. As we see in Fig. 23, in accordance with the convergence of the energy, the integrated distributions of the projected weight $\tilde{Q}(\tilde{\Omega})$ at any size show no distinct signal of insufficiency in sampling which is observed in the tail part of $\tilde{Q}(\tilde{\Omega})$ at $U/t = 15$ on 4×4 lattice as illustrated in Fig. 19.c. The existence of the cutoffs in the integrated distributions of Green's function $\tilde{Q}(\tilde{N})$ also supports the fact that the number of Monte Carlo steps 2.56×10^4 is sufficient in these systems (see Fig. 24).

Here, to confirm the convergence, we have calculated not only the ground state energy but also several physical quantities. First, we evaluate the equal-time spin and charge correlations defined by

$$S(\boldsymbol{k}) = \frac{1}{3N} \sum_{i,j} e^{i\boldsymbol{k} \cdot (\boldsymbol{r}_i - \boldsymbol{r}_j)} \langle \hat{\boldsymbol{S}}_i \cdot \hat{\boldsymbol{S}}_j \rangle, \tag{109}$$

$$N(\boldsymbol{k}) = \frac{1}{N} \sum_{i,j} e^{i\boldsymbol{k} \cdot (\boldsymbol{r}_i - \boldsymbol{r}_j)} \langle (\hat{n}_{i\uparrow} + \hat{n}_{i\downarrow})(\hat{n}_{j\uparrow} + \hat{n}_{j\downarrow}) \rangle. \tag{110}$$

As we see in Fig. 25, for all the system sizes, the peak values of the spin correlation $S(\pi, \pi)$ obtained by the PR-GBMC method are consistent with the values obtained by the AFQMC method. The peak values of $S(\pi, \pi)$ as well as the charge correlation $N(\pi, \pi)$ are listed in Table III. A small discrepancies observed in $N(\pi, \pi)$ may be attributed either to the effect of finite intervals in the imaginary time step of AFQMC or to the slight insufficiency of τ taken for the ground-state average in the PR-GBMC calculation.

Next, we show the superconducting correlation defined by

$$S_\alpha = \frac{1}{4} \sum_r P_\alpha(r), \tag{111}$$

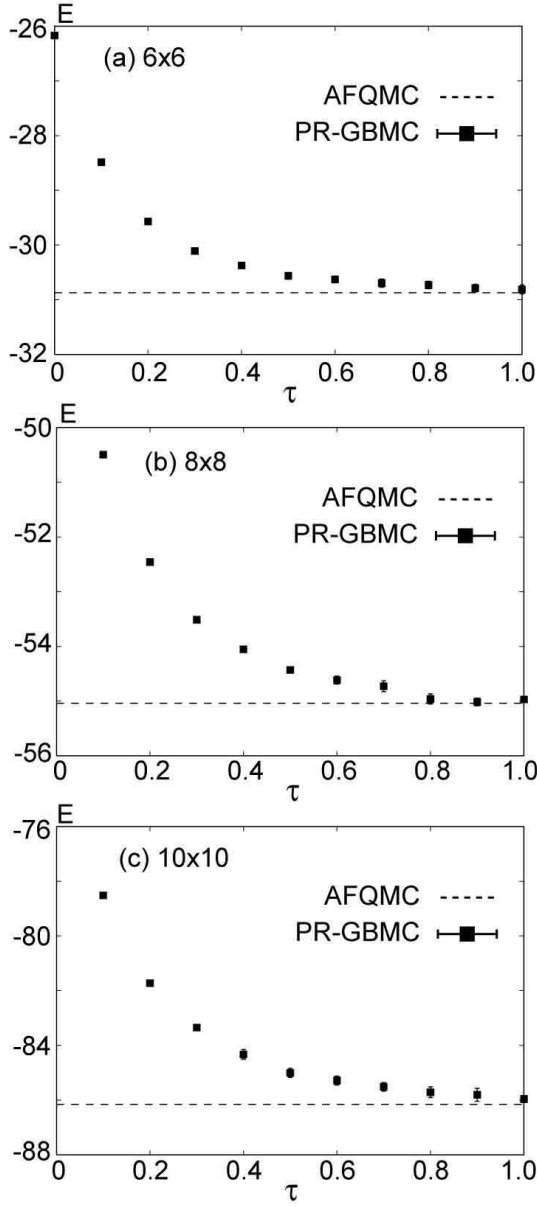where $P_\alpha(r)$ denotes the equal-time pairing correlation

Fig. 22.   Energy as functions of τ at $U/t = 4$ and $n = 1$ for (a) 6×6, (b) 8×8 and (c) 10×10 lattices under the periodic boundary condition. Dashed lines are the ground state energy obtained by the AFQMC method. Error bars are as large as the symbol size.
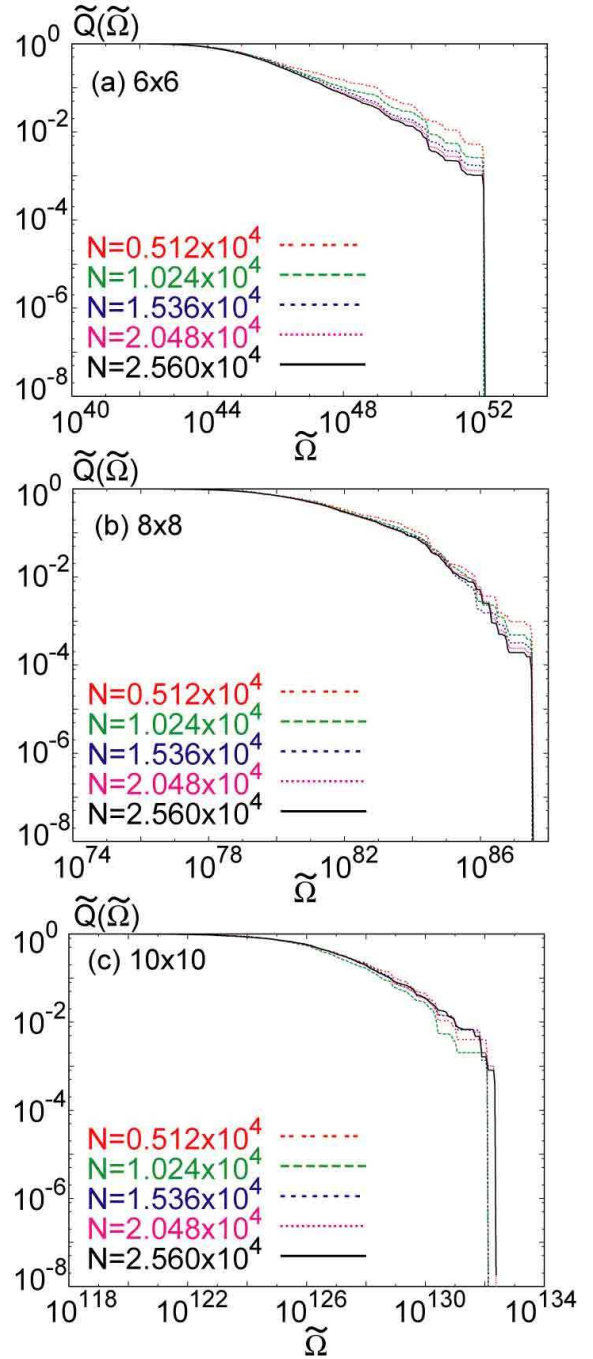

Fig. 23.   (color online): Integrated distribution of projected weight $\widetilde{\Omega}$ obtained at $U/t = 4$ $n = 1$ and $\tau = 1$ on (a) 6×6, (b) 8×8 and (c) 10×10 lattices under the periodic boundary condition. In all the panels, red (light), green (long dashed), blue (dashed), pink (dotted) and black (dark) curves represent the distribution obtained by 0.512×10^4, 1.024×10^4, 1.536×10^4, 2.048×10^4 and 2.560×10^4 Monte Carlo steps, respectively.

defined as

$$P_\alpha(r) = \frac{1}{2N} \sum_{i=1}^{N} \langle \Delta_\alpha^\dagger(i)\Delta_\alpha(i+r) + \Delta_\alpha(i)\Delta_\alpha^\dagger(i+r) \rangle,$$

(112)

where r is the distance from the i-th site and Δ_α is the superconducting order parameter. The latter is defined as

$$\Delta_\alpha(i) = \frac{1}{\sqrt{2}} \sum_r f_\alpha(r)(\hat{c}_{i\uparrow}\hat{c}_{i+r\downarrow} - \hat{c}_{i\downarrow}\hat{c}_{i+r\uparrow}),$$

(113)

where $f_\alpha(r)$ is the form factor of the pairing correlation defined as

$$f_{1s}(r) = 4\delta_{r_x,0}\delta_{r_y,0},$$

(114)

$$f_{2s}(r) = \delta_{r_y,0}(\delta_{r_x,1} + \delta_{r_x,-1})$$

$$+ \delta_{r_x,0}(\delta_{r_y,1} + \delta_{r_y,-1}),$$

(115)

$$f_{2d}(r) = \delta_{r_y,0}(\delta_{r_x,1} + \delta_{r_x,-1})$$

$$- \delta_{r_x,0}(\delta_{r_y,1} + \delta_{r_y,-1}),$$

(116)

$$f_{3s}(r) = \delta_{r_x,1}(\delta_{r_y,1} + \delta_{r_y,-1})$$

$$+ \delta_{r_x,-1}(\delta_{r_y,-1} + \delta_{r_y,1}),$$

(117)

$$f_{3d}(r) = \delta_{r_x,1}(\delta_{r_y,1} - \delta_{r_y,-1})$$
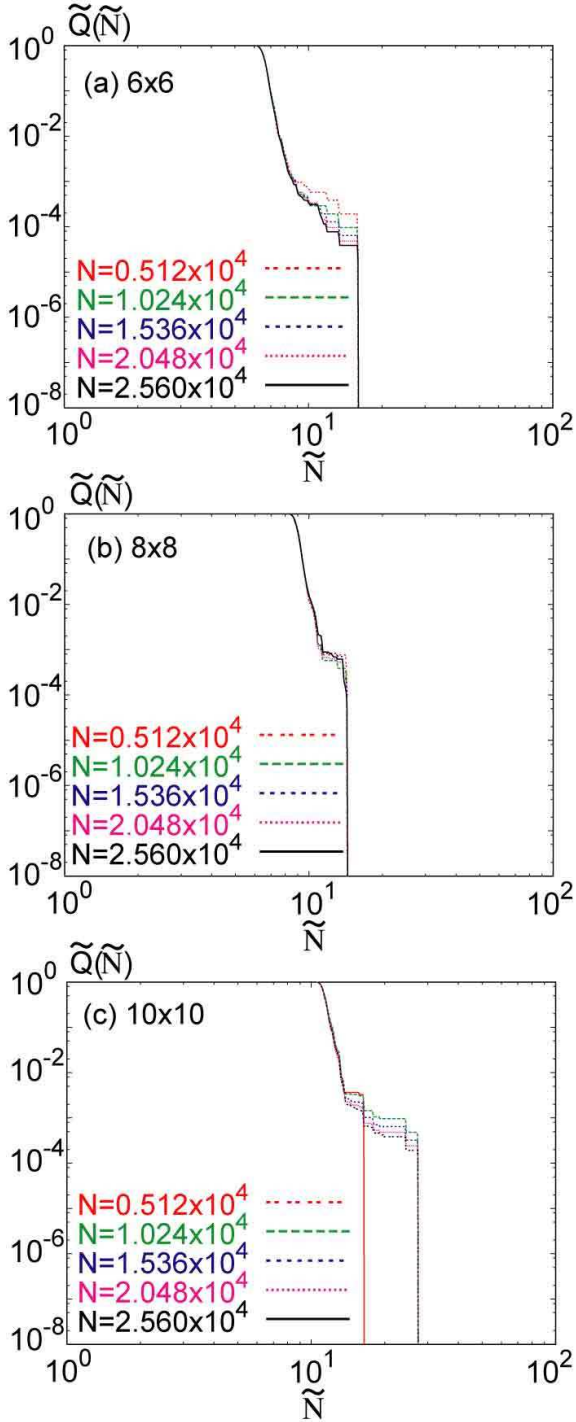
Fig. 24. (color online): Integrated distribution of Green's function $\tilde{N}$ obtained at $U/t = 4$ and $\tau = 1$ on (a) 6×6, (b) 8×8 and (c) 10×10 lattices under the periodic boundary condition. In all the panels, red (light), green (long dashed), blue (dashed), pink (dotted) and black (dark) curves represent the distribution obtained by 0.512×10^4, 1.024×10^4, 1.536×10^4, 2.048×10^4 and 2.560×10^4 Monte Carlo steps, respectively.

$$+ \delta_{r_x,-1}(\delta_{r_y,-1} - \delta_{r_y,1}), \qquad (118)$$

where δ_{ij} is Cronecker's delta. The suffices $\alpha = 1s, 2s, 2d, 3s$ and $3d$ represent the on-site s-wave, the extended s-wave directing along the x and y axes, the $d_{x^2-y^2}$-wave, the extended s-wave along the diagonals and the d_{xy}-wave, respectively. Numerical results for all
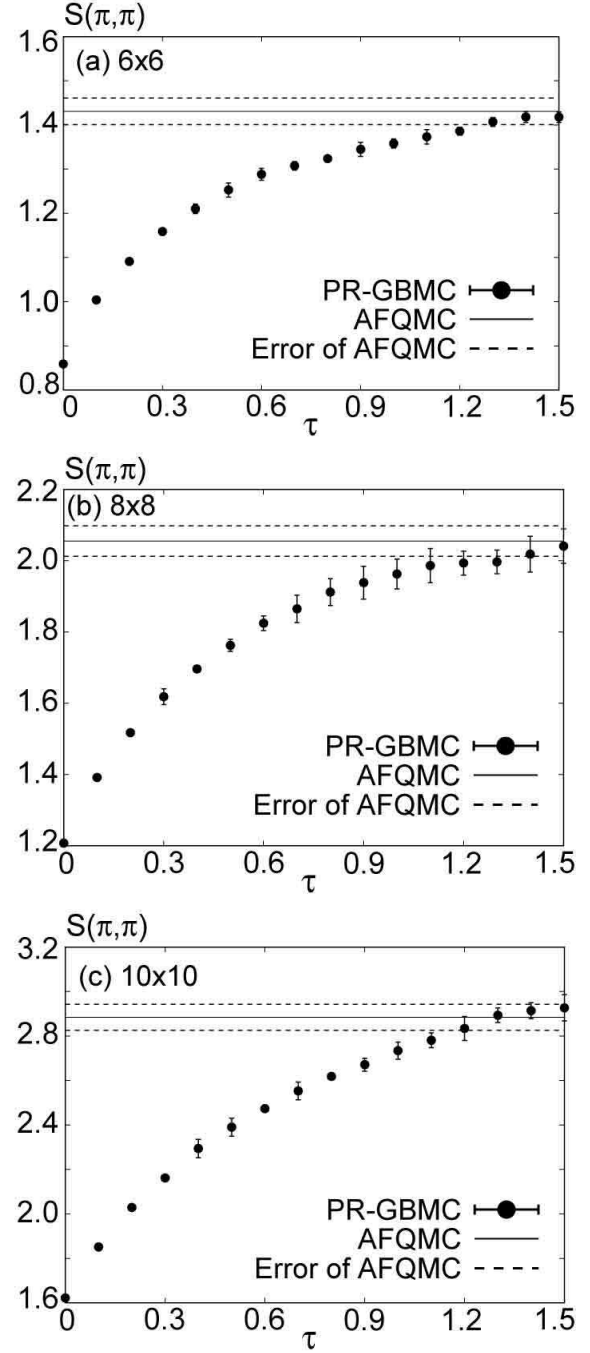

Fig. 25. Peak values of spin correlation $S(\pi, \pi)$ on (a) 6×6, (b) 8×8 and (c) 10×10 lattices under the periodic boundary condition at $U/t = 4$ and $n = 1$. In all the panels, solid and dashed lines represent the results obtained by the AFQMC method and their error ranges, respectively.

the quantities defined above are listed in Table III. Here, for comparison with the other method, we demonstrate the numerical results obtained by the AFQMC method with the Trotter discretization of $\Delta \tau = 0.025$ and the number of Monte Carlo steps $N_{\mathrm{mcs}} = 10^4$.

We have also calculated the momentum distribution defined by

$$n(\boldsymbol{k}) = \frac{1}{2N} \sum_{i,j} e^{i\boldsymbol{k} \cdot (\boldsymbol{r}_i - \boldsymbol{r}_j)} \langle \hat{c}_{i\uparrow}^\dagger \hat{c}_{j\uparrow} + \hat{c}_{i\downarrow}^\dagger \hat{c}_{j\downarrow} \rangle. \qquad (119)$$

In Table IV, we show the numerical results for the mo-

Table III. Comparison between PR-GBMC and AFQMC methods. For all the parameter sets, we employ the quantum-number projections onto the total momentum $\boldsymbol{K} = 0$, total z-component of the spin $S_z = 0$ and the total spin $S = 0$. All the simulation results are obtained by 2.56×10^4 Monte Carlo steps under the periodic boundary condition.

$6 \times 6, U/t = 4, n = 1$	PR-GBMC	AFQMC
Energy	-30.87 ± 0.02	-30.87 ± 0.03
$S(\pi, \pi)$	1.43 ± 0.01	1.43 ± 0.03
$N(\pi, \pi)$	0.409 ± 0.001	0.403 ± 0.001
S_{1s}	1.636 ± 0.005	1.610 ± 0.004
S_{2s}	1.188 ± 0.004	1.184 ± 0.001
S_{2d}	1.101 ± 0.007	1.116 ± 0.018
S_{3s}	0.539 ± 0.004	0.532 ± 0.003
S_{3s}	0.397 ± 0.003	0.399 ± 0.012

$8 \times 8, U/t = 4, n = 1$	PR-GBMC	AFQMC
Energy	-55.01 ± 0.03	-55.09 ± 0.06
$S(\pi, \pi)$	2.08 ± 0.04	2.05 ± 0.04
$N(\pi, \pi)$	0.420 ± 0.003	0.412 ± 0.002
S_{1s}	1.681 ± 0.011	1.650 ± 0.001
S_{2s}	1.191 ± 0.001	1.190 ± 0.001
S_{2d}	1.077 ± 0.004	1.110 ± 0.018
S_{3s}	0.515 ± 0.004	0.505 ± 0.002
S_{3s}	0.457 ± 0.005	0.457 ± 0.007

$10 \times 10, U/t = 4, n = 1$	PR-GBMC	AFQMC
Energy	-86.25 ± 0.10	-86.11 ± 0.03
$S(\pi, \pi)$	2.88 ± 0.03	2.88 ± 0.06
$N(\pi, \pi)$	0.424 ± 0.002	0.417 ± 0.001
S_{1s}	1.694 ± 0.006	1.668 ± 0.004
S_{2s}	1.195 ± 0.005	1.192 ± 0.001
S_{2d}	1.076 ± 0.005	1.111 ± 0.018
S_{3s}	0.510 ± 0.003	0.507 ± 0.002
S_{3s}	0.471 ± 0.005	0.459 ± 0.027

Table IV. Momentum distribution $n(k_x, k_y)$ obtained by PR-GBMC and AFQMC method. The results have been obtained from the same simulation as those of Table III

$6 \times 6, U/t = 4, n = 1$		
(k_x, k_y)	PR-GBMC	AFQMC
$(0, 0)$	0.967 ± 0.001	0.966 ± 0.001
$(\pi/3, \pi/3)$	0.928 ± 0.001	0.926 ± 0.001
$(2\pi/3, 2\pi/3)$	0.072 ± 0.001	0.074 ± 0.001
(π, π)	0.033 ± 0.001	0.034 ± 0.001

$8 \times 8, U/t = 4, n = 1$		
(k_x, k_y)	PR-GBMC	AFQMC
$(0, 0)$	0.967 ± 0.001	0.966 ± 0.001
$(\pi/4, \pi/4)$	0.951 ± 0.001	0.950 ± 0.001
$(\pi/2, \pi/2)$	0.500 ± 0.000	0.500 ± 0.000
$(3\pi/4, 3\pi/4)$	0.049 ± 0.001	0.050 ± 0.001
(π, π)	0.033 ± 0.001	0.034 ± 0.001

$10 \times 10, U/t = 4, n = 1$		
(k_x, k_y)	PR-GBMC	AFQMC
$(0, 0)$	0.968 ± 0.001	0.966 ± 0.001
$(\pi/5, \pi/5)$	0.959 ± 0.001	0.957 ± 0.001
$(2\pi/5, 2\pi/5)$	0.883 ± 0.003	0.878 ± 0.001
$(3\pi/5, 3\pi/5)$	0.117 ± 0.003	0.121 ± 0.001
$(4\pi/5, 4\pi/5)$	0.041 ± 0.001	0.043 ± 0.001
(π, π)	0.032 ± 0.001	0.033 ± 0.001

mentum distribution along the line from $(0, 0)$ to (π, π) in the momentum space.

In the numerical results shown in Tables III and IV, all the numerical data are essentially consistent each other. We also show a hard test of the convergence for the square-lattice Hubbard model with the next near-

est neighbor transfer $t' = 0.5t$ in one of the diagonal direction. This is nothing but the anisotropic triangular lattice. This geometrically frustrated lattice structure generates a serious difficulty in various simulations. The PIRG method offers the only available results.[8,9] Here PR-GBMC results for the 6×6 lattice with the periodic boundary condition at $U/t = 4$ shows the ground state enrgy $E = -32.54 \pm 0.05$ and the double occupancy $\langle D \rangle = 0.164 \pm 0.001$, which is favorably compared with the PIRG result of $E = -32.6 \pm 0.1$ and $\langle D \rangle = 0.168 \pm 0.006$. This parameter value is, as is estimated from the systematic studies of the PIRG result, just near the Mott transition point and provides us with a severe numerical benchmark. From the analysis of the size dependence, we conclude that the convergence and the applicability of the PR-GBMC method is not restricted by the system size.

5. Summary and Discussion

In this paper, we have reexamined the Gaussian-basis Monte Carlo method (GBMC) proposed by Corney and Drummond.[12,13] This method does not suffer from the minus sign problem for any Hamiltonian composed of up to two-body interactions (see Appendix A and B). However, the original method often shows systematic errors especially in the low-temperature region. We have elucidated how the systematic errors come from the slow relaxation caused by the trap in the excited states.

To overcome the systematic error, we have improved the quantum-number projection scheme proposed by Assaad et al.[16,17] to make it possible to combine the projection procedure in conjunction with the importance sampling of the original GBMC method. This method allows us to project out the excited states and improve the behavior of the probability distributions, which makes it possible to widen the region of tractable parameters in the GBMC method.

We have also discussed the applicability of our algorithm. In the PR-GBMC method, the convergence with the ground state becomes slower with the increase of the interaction strength U/t. This slow convergence is caused by the increasing inefficiency in the importance sampling procedure owing to the barrier in the phase-space coming from the strong on-site repulsion. Nevertheless, good convergence at $U/t = 10$ indicates a better efficiency of the PR-GBMC method compared to other numerical methods such as the AFQMC method. The system size dependence up to 10×10 lattice shows no distinct symptom of the slow convergence in physical quantities with the increase of the system size. Thus the slow convergence does not restrict the applicability of this method for large system sizes at least up to 10×10 lattices.

In the large U/t region, the slow convergence is tightly associated with the exponential broadening of the distribution of the weight Ω, which causes an expansion of the phase-space to be sampled and results in inefficient importance sampling. Especially, the exponential growth of Ω and corresponding expansion of the phase-space often results in the lack of rare event with large Ω. The lack of Monte Carlo samples with large Ω causes a spike structure in the tail part of the distribution $\tilde{P}(\tilde{\Omega})$ and

causes a plateau structure in the tail part of the integrated distribution $\tilde{Q}(\tilde{\Omega})$. Despite samples with large Ω contribute to physical quantities, such samples seldom appear at large U/t, which causes large statistical errors. Thus, with the increase of U/t, the convergence of the energy becomes slower accompanied by the increase of the statistical errors. If that is the case, the computation time required for the convergence of the energy with required statistical errors goes beyond allowed computation time, which determines the practical limitation of the PR-GBMC method. Therefore, it is advisable to monitor the convergence of the distribution to evaluate the number of Monte Carlo steps needed for the convergence, especially when one calculates large U/t regions.

With the inspection of the large U/t systems and the system size dependence, the PR-GBMC method offers a powerful tool which can be applied to the systems in several cases beyond tractable parameters of the conventional numerical methods such as the AFQMC and PIRG methods and is at least complementary to the existing methods.

Acknowledgements

Appendix A: Langevin Equations for a General Hamiltonian

In this Appendix, we show how to construct a Gaussian representation for a general Hamiltonian. Here for simplicity, we treat a general number-conserving Hamiltonian given by

$$\hat{H} = \sum_{ij\sigma} t_{ij\sigma}\hat{n}_{(i\sigma),(j\sigma)} + \sum_{ijkl\sigma\sigma'} V_{kl\sigma'}^{ij\sigma}\hat{n}_{(i\sigma),(j\sigma)}\hat{n}_{(k\sigma'),(l\sigma')}$$

$$= \sum_{ij\sigma}\hat{H}_{ij\sigma} + \sum_{ijkl\sigma\sigma'}\hat{H}_{ijkl\sigma\sigma'}. \quad (A\cdot1)$$

Using the operator identities in Eqs.(29)-(31), one-body term $\hat{H}_{ij\sigma}$ gives a contribution $A_{ij\sigma}^{\Omega}$ to the equation of Ω and a contribution $A_{ij\sigma}^{n_{xy}}$ to the drift term of the Green's function:

$$d\Omega = \sum_{ij\sigma} A_{ij\sigma}^{\Omega}\Omega d\tau, \quad dn_{xy} = \sum_{ij\sigma} A_{ij\sigma}^{n_{xy}} d\tau, \quad (A\cdot2)$$

where x (and y) denotes the site and the spin, i.e, $x = (i,\sigma)$. The concrete expressions of $A_{ij\sigma}^{\Omega}$ and $A_{ij\sigma}^{n_{xy}}$ are

$$A_{ij\sigma}^{\Omega} = -t_{ij\sigma}n_{(i\sigma),(j\sigma)}, \quad (A\cdot3)$$

$$A_{ij\sigma}^{n_{xy}} = -\frac{t_{ij\sigma}}{2}\left[n_{x,(j\sigma)}\{\delta_{(i\sigma),y} - n_{(i\sigma),y}\} + \{\delta_{x,(j\sigma)} - n_{x,(j\sigma)}\}n_{(i\sigma),y}\right]. \quad (A\cdot4)$$

Next, to guarantee the positive diffusion, we change the two-body term $\hat{H}_{ijkl\sigma\sigma'}$ by adding the identity $\hat{n}_{(i\sigma),(j\sigma)}^2 - \delta_{ij}\hat{n}_{(i\sigma),(j\sigma)} = 0$ as

$$\hat{H}_{ijkl\sigma\sigma'} = V_{kl\sigma'}^{ij\sigma}\hat{n}_{(i\sigma),(j\sigma)}\hat{n}_{(k\sigma'),(l\sigma')}$$

$$= -\frac{|V_{kl\sigma'}^{ij\sigma}|}{2}\left[\hat{n}_{(i\sigma),(j\sigma)} - \text{sign}(V_{kl\sigma'}^{ij\sigma})\hat{n}_{(k\sigma'),(l\sigma')}\right]^2$$

$$+ \frac{|V_{kl\sigma'}^{ij\sigma}|}{2}\left[\delta_{ij}\hat{n}_{(i\sigma),(j\sigma)} + \delta_{kl}\hat{n}_{(k\sigma'),(l\sigma')}\right], \quad (A\cdot5)$$

The two-body term then gives a contribution $A_{ijkl\sigma\sigma'}^{\Omega}$ to the equation of Ω, a contribution $A_{ijkl\sigma\sigma'}^{n_{xy}}$ to the drift term of n_{xy} and the contributions $B_{ijkl\sigma\sigma'}$, $C_{ijkl\sigma\sigma'}$ to the diffusion term of n_{xy}:

$$d\Omega = \sum_{ijkl\sigma\sigma'} A_{ijkl\sigma\sigma'}^{\Omega}\Omega d\tau, \quad (A\cdot6)$$

$$dn_{xy} = \sum_{ijkl\sigma\sigma'}\left[A_{ijkl\sigma\sigma'}^{n_{xy}}d\tau + B_{ijkl\sigma\sigma'}^{n_{xy}}dW_{ijkl\sigma\sigma'}^{(1)}\right.$$

$$\left. + C_{ijkl\sigma\sigma'}^{n_{xy}}dW_{ijkl\sigma\sigma'}^{(2)}\right], \quad (A\cdot7)$$

where the Wiener increment dW satisfies

$$\langle dW_{ijkl\sigma\sigma'}^{(r)}dW_{i'j'k'l'\eta\eta'}^{(r')}\rangle = d\tau\delta_{rr'}\delta_{ii'}\delta_{jj'}\delta_{kk'}\delta_{ll'}\delta_{\sigma\eta}\delta_{\sigma'\eta'}. \quad (A\cdot8)$$

The concrete expressions of each term become

$$A_{ijkl\sigma\sigma'}^{\Omega} = -\frac{V_{kl\sigma'}^{ij\sigma}}{2}\left[2\left\{n_{(i\sigma),(j\sigma)}n_{(k\sigma'),(l\sigma')}\right.\right.$$

$$\left. - n_{(i\sigma),(l\sigma')}n_{(k\sigma'),(j\sigma)}\right\}$$

$$\left. + n_{(i\sigma),(l\sigma')}\delta_{kj}\delta_{\sigma\sigma'} + n_{(k\sigma'),(j\sigma)}\delta_{il}\delta_{\sigma\sigma'}\right], \quad (A\cdot9)$$

$$A_{ijkl\sigma\sigma'}^{n_{xy}} = -\frac{V_{kl\sigma'}^{ij\sigma}}{2}\left[n_{(i\sigma),(j\sigma)}\left\{n_{x,(l\sigma')}(\delta_{(k\sigma'),y} - n_{(k\sigma'),y})\right.\right.$$

$$\left. + (\delta_{x,(l\sigma')} - n_{x,(l\sigma')})n_{(k\sigma'),y}\right\}$$

$$+ n_{(k\sigma'),(l\sigma')}\left\{n_{x,(j\sigma)}(\delta_{(i\sigma),y} - n_{(i\sigma),y})\right.$$

$$\left. + (\delta_{x,(j\sigma)} - n_{x,(j\sigma)})n_{(i\sigma),y}\right\}$$

$$- \left\{n_{(i\sigma),(l\sigma')} - \frac{1}{2}\delta_{(i\sigma),(l\sigma')}\right\}$$

$$\times \left\{n_{x,(j\sigma)}(\delta_{(k\sigma'),y} - n_{(k\sigma'),y})\right.$$

$$\left. + (\delta_{x,(j\sigma)} - n_{x,(j\sigma)})n_{(k\sigma'),y}\right\}$$

$$- \left\{n_{(k\sigma'),(j\sigma)} - \frac{1}{2}\delta_{(k\sigma'),(j\sigma)}\right\}$$

$$\times \left\{n_{x,(l\sigma')}(\delta_{(i\sigma),y} - n_{(i\sigma),y})\right.$$

$$\left.\left. + (\delta_{x,(l\sigma')} - n_{x,(l\sigma')})n_{(i\sigma),y}\right\}\right], \quad (A\cdot10)$$

and

$$B_{ijkl\sigma\sigma'}^{n_{xy}} = \sqrt{\frac{|V_{kl\sigma'}^{ij\sigma}|}{2}}\left[n_{x,(j\sigma)}\{\delta_{(i\sigma),y} - n_{(i\sigma),y}\}\right.$$

$$\left. - \text{sign}(V_{kl\sigma'}^{ij\sigma})n_{x,(l\sigma')}\{\delta_{(k\sigma'),y} - n_{(k\sigma'),y}\}\right], \quad (A\cdot11)$$

$$C_{ijkl\sigma\sigma'}^{n_{xy}} = \sqrt{\frac{|V_{kl\sigma'}^{ij\sigma}|}{2}} \left[\{\delta_{x,(j\sigma)} - n_{x,(j\sigma)}\} n_{(i\sigma),y} \right.$$

$$\left. - \mathrm{sign}(V_{kl\sigma'}^{ij\sigma}) \{\delta_{x,(l\sigma')} - n_{x,(l\sigma')}\} n_{(k\sigma'),y} \right]. \tag{A·12}$$

In all, the general Hamiltonian (A·1) gives the Langevin equations of Ito-type

$$d\Omega = \left[\sum_{ij\sigma} A_{ij\sigma}^{\Omega} + \sum_{ijkl\sigma\sigma'} A_{ijkl\sigma\sigma'}^{\Omega} \right] \Omega d\tau, \tag{A·13}$$

$$dn_{xy} = \left[\sum_{ij\sigma} A_{ij\sigma}^{n_{xy}} + \sum_{ijkl\sigma\sigma'} A_{ijkl\sigma\sigma'}^{n_{xy}} \right] d\tau$$

$$+ \sum_{ijkl\sigma\sigma'} \left[B_{ijkl\sigma\sigma'}^{n_{xy}} dW_{ijkl\sigma\sigma'}^{(1)} + C_{ijkl\sigma\sigma'}^{n_{xy}} dW_{ijkl\sigma\sigma'}^{(2)} \right]. \tag{A·14}$$

Appendix B: Completeness and Positivity of the Gaussian Basis

In order to establish a phase-space representation based on the Gaussian operators, one must show that they form a complete basis for the class of density-matrix operators which we wish to represent. Furthermore, the expansion of any density-matrix operators in terms of the Gaussian operators must involve the positive coefficients. For the self-contained description of this paper, we prove this 'positive completeness' by following the idea of J. F. Corney.[24]

To prove this 'positive completeness', we will relate the Gaussian operators to the number-state projection operators which form a complete basis set and we will show that any term which appears in a number-state expansion of a density-matrix operator can be written as a sum of Gaussian operators with positive coefficients. Here, we prove only the number-conserving case in this Appendix, a proof for the most general Gaussian can be similarly constructed.[24] To this end, we rewrite the Gaussian operator in terms of $\boldsymbol{\mu}$ which is defined as

$$\boldsymbol{\mu} = 2\boldsymbol{I} + (\boldsymbol{n}^T - \boldsymbol{I})^{-1}. \tag{B·1}$$

A number-conserving Gaussian operator $\hat{\Lambda}(\boldsymbol{n})$ is then written as

$$\hat{\Lambda} = \det(\boldsymbol{I} - \boldsymbol{n}) : e^{-\hat{\boldsymbol{b}}^{\dagger}[2\boldsymbol{I}+(\boldsymbol{n}^T-\boldsymbol{I})^{-1}]\hat{\boldsymbol{b}}} :$$

$$= \frac{1}{\det(2\boldsymbol{I} - \boldsymbol{\mu})} : e^{-\hat{\boldsymbol{b}}^{\dagger}\boldsymbol{\mu}\hat{\boldsymbol{b}}} :$$

$$=: \prod_{i,j=1}^{M} (\hat{1} - \hat{b}_i^{\dagger}\mu_{ij}\hat{b}_j) : / \det(2\boldsymbol{I} - \boldsymbol{\mu}). \tag{B·2}$$

To avoid singular behavior, the limit of any $n_{jj} \to 1$ is taken only in the normalized form of the Gaussian.

B.1   Number-state expansion

Let $\vec{n}$ be a Fermionic occupation number vector $\vec{n} = (n_1, n_2, \cdots, n_M)$, where $n_i = 0, 1$, then a complete set of Fermionic number-state is represented by $\{|\vec{n}\rangle\}$, where

$\{|\vec{n}\rangle\}$ runs over all the 2^M permutations. This set defines a complete operator basis of dimension 2^{2M}, and it enables us to expand the density-matrix operator as

$$\hat{\rho} = \sum_{\vec{n}} \sum_{\vec{m}} |\vec{n}\rangle\langle\vec{n}|\hat{\rho}|\vec{m}\rangle\langle\vec{m}|$$

$$= \sum_{\vec{n}} \sum_{\vec{m}} \rho_{\vec{n}\vec{m}} |\vec{n}\rangle\langle\vec{m}| = \sum_{\vec{n}} \sum_{\vec{m}} \hat{\rho}_{\vec{n}\vec{m}}, \tag{B·3}$$

where we impose a number-conserving condition $\sum_i n_i = \sum_i m_i$. From the positive definiteness of the density-matrix operator, all the diagonal density-matrix elements are real and positive: $\rho_{\vec{n}\vec{n}} \geq 0$. Here we require additionally that $\rho_{\vec{n}\vec{n}} \leq 1$ for the normalization.

Since a density-matrix operator is Hermitian, it can be always diagonalized. Let $|\Psi_k\rangle = \sum_{\vec{n}} C_{k\vec{n}} |\vec{n}\rangle$ are the eigenvectors and P_k the corresponding positive eigenvalues of the density matrix, we can write

$$\hat{\rho} = \sum_k P_k |\Psi_k\rangle\langle\Psi_k| = \sum_k \sum_{\vec{n},\vec{m}} P_k C_{k\vec{n}} C_{k\vec{m}}^* |\vec{n}\rangle\langle\vec{m}|. \tag{B·4}$$

Thus the coefficients of the number-state expansion can be represented as

$$\rho_{\vec{n}\vec{m}} = \sum_k P_k C_{k\vec{n}} C_{k\vec{m}}^* = \sum_k D_{k\vec{n}} D_{k\vec{m}}^*, \tag{B·5}$$

where $D_{k\vec{n}} = \sqrt{P_k} C_{k\vec{n}}$. By using a Cauchy-Schwartz inequality, the magnitude of these coefficients is given by

$$|\rho_{\vec{n}\vec{m}}|^2 = \sum_{k,k'} D_{k,\vec{n}} D_{k\vec{m}}^* D_{k'\vec{n}}^* D_{k'\vec{m}}$$

$$\leq \sum_{k,k'} D_{k,\vec{n}} D_{k\vec{n}}^* D_{k'\vec{m}} D_{k'\vec{m}}^*$$

$$= \rho_{\vec{n}\vec{n}} \rho_{\vec{m}\vec{m}} \leq 1. \tag{B·6}$$

Thus the magnitude of any off-diagonal element is bounded, Conversely, any diagonal element is at least as large as the squared magnitude of anything else on the same row or column:

$$|\rho_{\vec{n}\vec{m}}|^2 \leq \rho_{\vec{n}\vec{n}} \rho_{\vec{m}\vec{m}} \leq \rho_{\vec{n}\vec{n}} \tag{B·7}$$

$$|\rho_{\vec{m}\vec{n}}|^2 \leq \rho_{\vec{m}\vec{m}} \rho_{\vec{n}\vec{n}} \leq \rho_{\vec{n}\vec{n}}. \tag{B·8}$$

Thus we obtain the lower limit of the diagonal elements:

$$\rho_{\vec{n}\vec{n}} \geq \max\left(|\rho_{\vec{n}\vec{m}}|^2, |\rho_{\vec{m}\vec{n}}|^2\right)$$

$$\geq \frac{1}{2(M-1)} \sum_{\vec{m}\neq\vec{n}} \left(|\rho_{\vec{n}\vec{m}}|^2 + |\rho_{\vec{m}\vec{n}}|^2\right)$$

$$\geq \frac{1}{2(2^M-1)} \sum_{\vec{m}\neq\vec{n}} \left(|\rho_{\vec{n}\vec{m}}|^2 + |\rho_{\vec{m}\vec{n}}|^2\right). \tag{B·9}$$

The number-state expansion of the density-matrix operator can then be written as

$$\hat{\rho} = \sum_{\vec{n}} \Delta\rho_{\vec{n}\vec{n}} |\vec{n}\rangle\langle\vec{n}| + \sum_{\vec{n}} \sum_{\vec{m}\neq\vec{n}} \frac{1}{2(2^M-1)} |\rho_{\vec{n}\vec{m}}|^2 \hat{P}_{\vec{n}\vec{m}}(\rho),$$

$$\tag{B·10}$$

where

$$\hat{P}_{\vec{n}\vec{m}}(\rho) = |\vec{n}\rangle\langle\vec{n}| + |\vec{m}\rangle\langle\vec{m}| + \frac{2(2^M - 1)}{\rho_{\vec{n}\vec{m}}^*}|\vec{n}\rangle\langle\vec{m}| \quad \text{(B·11)}$$

$$\Delta\rho_{\vec{n}\vec{n}} = \rho_{\vec{n}\vec{n}} - \frac{1}{2(2^M - 1)}\sum_{\vec{m}\neq\vec{n}}\left(|\rho_{\vec{n}\vec{m}}|^2 + |\rho_{\vec{m}\vec{n}}|^2\right) \geq 0. \quad \text{(B·12)}$$

Since all the coefficients of the new expansion (B·10) are positive, it is sufficient to prove that each operator in the expansion (B·10) can be written as a Gaussian or as a positive sum over Gaussians.

B.2 Diagonal number-state projector

First we show that the diagonal number-state projector $|\vec{n}\rangle\langle\vec{n}|$ in the new expansion (B·10) with positive coefficients $\Delta\rho_{\vec{n}\vec{n}}$ corresponds to a Gaussian operator. For individual ladder operators, one has the well-known identities:

$$\hat{b}_j = |0\rangle_j\langle 1|_j, \quad \hat{b}_j^\dagger = |1\rangle_j\langle 0|_j. \quad \text{(B·13)}$$

If we set $n_{ij} = n_j\delta_{ij}$ in Eq. (B·1), then $\mu_{ij} = \delta_{ij}(1 - 2n_j)/(1 - n_j)$ and the Gaussian operator reduces to

$$\hat{\Lambda} =: \prod_{i,j=1}^M (\hat{1} - \hat{b}_i^\dagger \mu_{ij}\hat{b}_j) : / \det(2\boldsymbol{I} - \boldsymbol{\mu})$$

$$= \prod_{j=1}^M \left[(1 - n_j)\hat{b}_j\hat{b}_j^\dagger + n_j\hat{b}_j^\dagger\hat{b}_j\right]$$

$$= \prod_{j=1}^M \left[(1 - n_j)|0\rangle_j\langle 0|_j + n_j|1\rangle_j\langle 1|_j\right]. \quad \text{(B·14)}$$

Thus, if n_j is chosen as 0 or as 1, the Gaussian operator itself can be regarded as a diagonal number-state projector:

$$|\vec{n}\rangle\langle\vec{n}| = \hat{\Lambda}(n_j\delta_{ij}). \quad \text{(B·15)}$$

B.3 Off-diagonal number-state projector

Second we show that the mixed projector $\hat{P}_{\vec{n}\vec{m}}(\rho)$ in the expansion (B·10) corresponds to a positive sum over Gaussians. Consider a Gaussian operator with

$$n_{ij} = n_j\delta_{ij} + \sum_{k=1}^{N_k}\nu_k\delta_{i,r(k)}\delta_{j,s(k)}, \quad \text{(B·16)}$$

where the n_j for each j is each either 0 or 1 and the locations $(r(k), s(k))$ of the N_k nonzero off-diagonal elements ν_k satisfy $r(k) \neq r(k'), s(k) \neq s(k')$ and $r(k) \neq s(k')$ for any $k, k'(k \neq k')$. In other words, if there is a nonzero element in the off-diagonal location (i, j), then there will be no other element in the i-th row and j-th column and

none in the j-th row and i-th column, i.e.,

$$\boldsymbol{n} = \begin{bmatrix} n_1 & & & 0 & & & 0 & & & \\ & \ddots & & \vdots & & & \vdots & & & \\ & & \ddots & 0 & & & 0 & & & \\ 0 & \cdots & 0 & n_i & 0 & \cdots & 0 & n_{ij} & 0 & \cdots & 0 \\ & & & 0 & \ddots & & \vdots & & & \\ & & & \vdots & & \ddots & 0 & & & \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & n_j & 0 & \cdots & 0 \\ & & & 0 & & & 0 & \ddots & & \\ & & & \vdots & & & \vdots & & \ddots & \\ & & & 0 & & & 0 & & & n_M \end{bmatrix}. \quad \text{(B·17)}$$

This structure means that $\det(\boldsymbol{I} - \boldsymbol{n}) = \prod_{j=1}^M (1 - n_j)$ and

$$\mu_{ij} = \delta_{ij}\frac{1 - 2n_j}{1 - n_j}$$

$$- \sum_{k=1}^{N_k}\delta_{j,r(k)}\delta_{i,s(k)}\frac{\nu_k}{(1 - n_{r(k)})(1 - n_{s(k)})}. \quad \text{(B·18)}$$

Here again, we take the limit $n_j \to 1$ only in the normalized form of the Gaussian to avoid the singularity. With these conditions, the Gaussian operator reduces to

$$\hat{\Lambda} =: \prod_{i,j=1}^M \left(1 - \hat{b}_i^\dagger \mu_{ij}\hat{b}_j\right) : / \det(2\boldsymbol{I} - \boldsymbol{\mu})$$

$$=: \prod_{j=1}^M \left[(1 - n_j)\hat{b}_j\hat{b}_j^\dagger + n_j\hat{b}_j^\dagger\hat{b}_j\right]$$

$$\times \prod_{k=1}^{N_k}\left[1 + \frac{\nu_k\hat{b}_{s(k)}^\dagger\hat{b}_{r(k)}}{(1 - n_{r(k)})(1 - n_{s(k)})}\right] :$$

$$=: \prod_{j=1}^M \left[(1 - n_j)|0\rangle_j\langle 0|_j + n_j|1\rangle_j\langle 1|_j\right]$$

$$\sum_K \prod_{k\in K}\frac{\nu_k\hat{b}_{s(k)}^\dagger\hat{b}_{r(k)}}{(1 - n_{r(k)})(1 - n_{s(k)})} :, \quad \text{(B·19)}$$

where the sum over K is the sum over all the possible subsets of $\{1, 2, \cdots, N_k\}$ and has 2^{N_k} terms. For the j-th mode, if there is a k such that $s(k) = j$, then the diagonal number projector for the j-th mode $|n_j\rangle\langle n_j|$ is replaced by the off-diagonal $|1\rangle_j\langle 0|_j$, or if $r(k) = j$, the conjugate projector $|0\rangle_j\langle 1|_j$ is created. Thus the Gaussian operator can be represented by a sum over number-state projectors:

$$\hat{\Lambda} = \sum_K(\pm)\prod_{k\in K}\nu_k\left|\vec{n}_{r(K)}^{s(K)}\right\rangle\left\langle\vec{n}_{s(K)}^{r(K)}\right| \equiv \hat{\Lambda}_{\boldsymbol{\nu}}(\vec{n}), \quad \text{(B·20)}$$

where the j-th element of the vector $\vec{n}_{s(K)}^{r(K)}$ is defined as

$$\left\{\vec{n}_{s(K)}^{r(K)}\right\}_j = \begin{cases} 1 & j = r(k), \exists k \in K \\ 0 & j = s(k), \exists k \in K \\ n_j & \text{otherwise} \end{cases}. \quad \text{(B·21)}$$

A minus sign appears if an odd number of transpositions are required to put all the annihilation and the creation

operators in a canonical order. In this sum over projectors, the diagonal projector $|\vec{n}\rangle\langle\vec{n}|$ is contained with coefficient 1. In the sum, there also exists the projector that transposes all the $2N_k$ specified modes with coefficient $\prod_{k=1}^{N_k} \nu_k$. The sum also contains projectors that transpose only subsets of these modes. In total, there are 2^{N_k} terms.

Next by adding other Gaussian operators, we eliminate all the intermediate terms from the sum in Eq. (B·20) and leave only $K = N_k$ terms. First we add the N_k Gaussians with one fewer off-diagonal element in the $\boldsymbol{n}$ matrix, to cancel the projectors that transpose $N_k - 1$ modes. Second we add $_{N_k}C_2$ Gaussians with two fewer off-diagonal elements, to cancel the projectors that transpose $N_k - 2$ modes. This process is repeated until all the $\sum_{k=1}^{N_k-1} {_{N_k}C_k} = 2^{N_k} - 2$ intermediate terms are removed. Finally, we obtain

$$\sum_{\boldsymbol{\nu}'\subseteq\boldsymbol{\nu}} \hat{\Lambda}_{\boldsymbol{\nu}'}(\vec{n}) = (2^{N_k} - 1)|\vec{n}\rangle\langle\vec{n}|$$

$$\pm \prod_{k=1}^{N_k} \nu_k \left| \vec{n}_{r(K)}^{s(K)} \right\rangle \left\langle \vec{n}_{s(K)}^{r(K)} \right|, \qquad \text{(B·22)}$$

where the sum indexed by subsets of $\boldsymbol{\nu}$ refers to the sum described above.

By adding $\hat{\Lambda}_{\boldsymbol{\nu}'}(\vec{n})$ and $\hat{\Lambda}_{\boldsymbol{\nu}'}(\vec{m})$, with different diagonal components $\vec{n} = \vec{n}_{r(K)}^{s(K)}$ and $\vec{m} = \vec{m}_{r(K)}^{s(K)}$, respectively, we obtain

$$\hat{P}_{\vec{n}\vec{m}}(\rho) = |\vec{n}\rangle\langle\vec{n}| + |\vec{m}\rangle\langle\vec{m}| + \frac{2(2^M - 1)}{\rho_{\vec{n}\vec{m}}^*}|\vec{n}\rangle\langle\vec{m}|$$

$$= \frac{1}{2^{N_k} - 1} \sum_{\boldsymbol{\nu}'\subseteq\boldsymbol{\nu}} \left[\hat{\Lambda}_{\boldsymbol{\nu}'}(\vec{n}) + \hat{\Lambda}_{\boldsymbol{\nu}'}(\vec{m}) \right], \quad \text{(B·23)}$$

where

$$\frac{1}{\rho_{\vec{n}\vec{m}}^*} = \pm \frac{1}{(2^M - 1)(2^{N_k} - 1)} \prod_k \nu_k. \qquad \text{(B·24)}$$

Thus it is proven that $\hat{P}_{\vec{n}\vec{m}}(\rho)$ can be represented by a positive sum over Gaussians and hence it is shown that any number-conserving density-matrix operator can be expanded by the Gaussian operators with positive coefficients.

1) R. Blankenbecler, D. J. Scalapino, and R. L. Sugar : Phys. Rev. D **24** (1981) 2278.
2) S. Sorella, S. Baroni, R. Car, and M. Parrinello : Eur. Phys. Lett. **8** (1989) 663.
3) M. Imada, and Y. Hatsugai : J. Phys. Soc. Jpn. **58** (1989) 3752.
4) N. Furukawa, and M. Imada : J. Phys. Soc. Jpn. **61** (1992) 3331.
5) S. R. White : Phys. Rev. B **48** (1993) 10345.
6) M. Imada, and T. Kashima : J. Phys. Soc. Jpn. **69** (2000) 2723.
7) T. Kashima, and M. Imada : J. Phys. Soc. Jpn. **70** (2001) 2287.
8) T. Kashima, and M. Imada : *ibid.* **70** (2001) 3052.
9) H. Morita, S. Watanabe, and M. Imada : J. Phys. Soc. Jpn. **71** (2002) 2109.
10) S. Watanabe, and M. Imada : J. Phys. Soc. Jpn. **73** (2004) 1251.
11) T. Mizusaki, and M. Imada : Phys. Rev. B **69** (2004) 125110.
12) J. F. Corney, and P. D. Drummond : Phys. Rev. B **73** (2006) 125112.
13) J. F. Corney, and P. D. Drummond : J. Phys. A: Math. Gen. **39** (2006) 269.
14) P. D. Drummond, and C. W. Gardiner : J. Phys. A **13** (1980) 2353.
15) J. F. Corney, and P. D. Drummond : Phys. Rev. A **68** (2003) 063822.
16) F. F. Assaad, P. Werner, P. Corboz, E. Gull, and M. Troyer : Phys. Rev. B **72** (2005) 224518.
17) F. F. Assaad, P. Werner, P. Corboz, E. Gull, and M. Troyer : *Effective Models for Low-Dimensional Strongly Correlated Systems* (American Institute of Physics, 2006), edited by G. G. Batrouni, and D. Poilblanc, p. 204.
18) C. W. Gardiner : *Handbook of Stochastic Methods* (Springer-Verlag, Berlin, 1983).
19) P. D. Drummond, and I. K. Mortimer : J. Comp. Phys. **93** (1991) 144.
20) P. Kloeden, E. Platen, and H. Schurz : *Numerical Solution of SDE Through Computer Experiments* (Springer-Verlag, Berlin, 1994).
21) M. R. Dowling, M. J. Davis, P. D. Drummond, and J. F. Corney : J. Comp. Phys. **220** (2007) 549.
22) M. C. Buonaura, and S. Sorella : Phys. Rev. B **57** (1998) 11446.
23) W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery : *NUMERICAL RECIPES in Fortran 77* (Cambridge University Press, 1992).
24) J. F. Corney : private communication.